Green Analytics

Finding Correlations between Deforestation and Socio-Economic Factors

Piyush Jain, Sahil Poonatar, Sanjay Naik, Saurabh Sinha, Subhabrata Chaudhuri, Suneet Taparia

Motivation

- Economic development comes at the cost of deforestation

- This relationship has been difficult to establish empirically due to
  - insufficient data
  - complex geo-climactic factors

Our Goal

- Determine the correlation between deforestation and various indexes of socio- economic development
    - Analyze each of these dimensions against deforestation data
    - Generate a consolidated score covering all Key Performing Indicators for every country
- Visualize the output in a user-friendly interactive dashboard

How is it done Today? Limitations?

- The Environmental Kuznets curve (EKC) hypothesizes the relationship between various indicators of environmental degradation

- Current studies on this topic are based on a single dimension and have not been inclusive of other factors

Our Approach

- Obtain available data from reliable sources (e.g. worldbank.org)
- Analyze the socio-economic factors that are most relevant for this study
- Model building based on this data using Python and SQLite
- An interactive visualization will be created with Tableau

How are we Distinguishing Ourselves?

- Include more dimensions of socio-economic development
- Bring meaningful comparisons under one hood
- User will be able to interact with the data

Who cares? What's its Impact?

- Government and environmental agencies can chart out policies for sustainable development

- Supporting tool for guiding environmental strategies

Risks

Not arriving at any conclusive correlation between deforestation levels and any of the socio-economic factors.

Payoffs

Be able to provide a platform for deeper analysis into this increasingly important subject.

Cost

- We will be using
  - readily available data from reputed and official sources
  - open-source technology

- We do not expect any major cost developing this project

How long will it take?

Prototype consisting of end-to-end data pipeline + model algorithm including interactive visualizations in **8 to 10 weeks**

How will Progress be Measured?

- Design specifications - 4 weeks
- Build and Test cycle(s) - 3 weeks
- Final updated design documentation -1 week
- Final presentation - 2 weeks

Progress will be measured using stage gate checklists