

CSE 490C, Programming Assignment 4, Due Tuesday, December 11, 2018, at 11:59 pm

Data Analysis of a DataSet

For this assignment, you will be understanding and processing a dataset available at [India Lights API](#). The dataset contains light output at night for 20 years, from 1993 to 2013, for 600,000 villages across India. This assignment has to be done individually. You will need to submit a .zip file, containing both your code and write-up, through Canvas and demo your assignment to TAs.

Data Visualization Specifications: For this part of the assignment, you'll explore the distribution of light output across India by creating time series graphs. Using the API, visualize light output (through bar graphs or line graphs) according to the three points below. Create a user interface to display the graphs; you might have three tabs that display data for each of the three points. Within each tab, add a user interface component (such as a slider or scroll bar) that will allow users to **view the same type of graph for each of the 20 years**.

- Graph the median of monthly light outputs for each state. Do this for each year 1993-2013.
- Split all districts into quintiles based on their median of monthly light outputs. Graph the average light output of each quintile. Indicate the percent change in average light output between the districts in the upper and lower quintile (e.g., this could be a sixth bar or a line indicating height difference with a percent change label). Do this for each year 1993-2013.
- Graph the percent change in median yearly light output between capital districts and the rest of the districts in a state. Do this for each year 1993-2013. A .csv file of Indian states and their capitals can be downloaded [here](#).

Please note a few things about the API and assignment requirements.

- Correction to the API: to get the list of districts that you can request time series data for, the command is `http://api.nightlights.io/districts`, *not* `http://api.nightlights.io/regions`.
- Some months do not have night light data, possibly because of cloud cover. You do not have to take these months into account for calculating descriptive statistics.

- Note that percent change is not the same as percent difference. The formula for percent change is $(V2 - V1)/V1 * 100\%$
- Follow best practices for creating graphs and interacting with data. Clearly label graph titles, axes, intervals, and the meaning of scroll bars/sliders.

Write-up Specifications and Screenshot Requirements: In around 500 words, describe what you learned about changes in the distribution of light output in India from visualizing this dataset. How has light output from each state changed over time? How have the disparities in light output changed over time? Describe what might explain the trends you saw in the data, based on research about electrification initiatives in India and economic activity in different states or rural vs. urban areas. Cite sources as necessary and point to specific differences over time in your data visualization to support your reasoning. Also describe limitations of the data, such as what it can and cannot tell you about electrification/power usage.

In the same document, after the written portion, should be a total of 15 screenshots: graphs from the years 1993, 1998, 2003, 2008, and 2013, for each of the three sets. Additionally, when you reference specific trends, you should include screenshots of graphs that best demonstrate those trends as well, if they are not covered in the set of 15.

Choice of Technology: Since the assignment submission is based on the analysis and presentation of a dataset rather than any implementation specifics, you can use the programming language and libraries of your choice for creating the interface and graphs.

Submission: You should submit the code, the write-up, and screenshots of graphs through Canvas in a .zip file (firstname.lastname.zip) by the due date. Make sure to add your name, student ID, and comments within the code to explain your implementation. If you include external APIs, header files or packages, make sure to briefly explain their need and use in the implementation.