

# SateLIFE Proposal

## Group Name

SateLIFE

## Group Members

Bethany Bailey, Cooper Nederhood, Joseph Denby, Laurence Warner

## Data Set

Google Earth Engine satellite images (as raster files that can be preprocessed and downloaded as numpy arrays) from 2000-2017.

## Data Set Size

We are currently unsure of the exact size of the data set in GB. We are planning to download it as a three dimensional matrix with multiple bands from the Google Earth Engine. However, we are confident that this data set will be sufficiently large for the project as it will be easily scaled up by choosing multiple cities or covering a larger area.

## Number of Records

### Cities

We're planning to look at 3-6 cities, starting with Brazzaville, Republic of the Congo and Kinshasa, Democratic of the Congo. These cities are the two closest city capitals in the world (other than the Vatican and Rome). We are interested in using these two cities as a case study to analyze recent development and environmental factors.

### Bands

Each city will have satellite bands for the following dimensions:

- Vegetation Index
- LandSat 7 RGB Image
- DMSP Nightlights
- MODIS Air Quality
- MODIS Surface Temperature

The resolution of these bands is 30m/pixel, with coverage approximately every two weeks. We will cover approximately 17 years (starting in 2000). Each band will be pre-processed in the Google Earth Engine and then downloaded as a raster file/numpy arrays for additional analysis. Each pixel will be represented by a cell in the matrix on 5 dimensions (the five bands mentioned above). Thus, our data will be a three dimensional matrix that is the size of the area we choose wide and long, and five units deep.

## Hypotheses

We chose our two initial cities (Brazzaville, Republic of the Congo and Kinshasa, Democratic of the Congo) because they are the two closest city capitals (other than the Vatican and Rome). We are interested in using these two cities as a case study to analyze recent development and environmental factors. We want to analyze how urban development influences environmental changes across time, such as air quality, vegetation, and temperature. We would like to look at this on a entire-city (macro) level and a sub-city (micro) level (such as the macro analysis of differences between the cities and the micro analysis of differences within the urban environment of each city). After we have analyzed these two cities, we would like to apply our analysis to other developing/developed cities.

## Ideas for scalable algorithms

Initially, we plan to use javascript in the Google Earth Engine Code Editor to gather and pre-process our data. Due to the size and complexity of our data, we are planning to look at each pixel value independently (without the context of the pixels around it), which will hopefully allow for parallelizable algorithms to process the data. Other than that, we don't have specific ideas for algorithms yet.