We will want to do the following for each disease (yellow fever as an example)

Determine the following data points in final DataFrame/Table:

Disease Name

Disease Year 1

Global avg. Temperature

Disease Max Lat

Disease Min Lat

Disease Year 2

Global avg. Temperature

Disease Max Lat

Disease Min Lat

… etc.

Should consider *database normalization* for the SQL tables to gain better practices with relational dbs, once made will also make querying much easier.

Create a visual of Disease max/min lat on top of a world map. Create a time slider with years of associated data.

A disease table will need to be created

Year , Country, Instances/Counts, Max Lat of the country with most counts

Intermediate table

Year, Max Lat

Then run regression for Lat vs. Temp

Use the API’s from the WHO website to extract data

We will use MySQL and store all the data

Determine the highest (in terms of latitude) incidents of each disease

Determine the country

Determine the Max/Min Lat

The query will need to be multi indexed - Database normalization as alternative

Year then the country and there will be multiple countries per year

**Normalization Task: 3NF**

Create the following tables:

1. Country ID, Country, latitude (Max Lat, Min Lat)( will do a web search maybe WHO site?)
2. Disease ID, Disease ( manual creation)
3. Year, Temperature( cdc global .gov scrape ...drop into pandas)
4. TBD: Year, DiseaseID, CountryID, Number of Infections(API calls to sites)

**Proposal**:

Question:

Does the expanding temperate zone (global warming) influence the occurrence of tropical diseases? These diseases historically have been isolated to specific latitudes.

**Approach**:

Using historical data we will extract temperatures, latitudes, infection counts to determine if there is any correlation.

**Potential visuals:**

* A world map that will have a selection box to allow selection of the diseases which we are searching for.
* A slider bar that will allow the user to view a specific time period.
* A regression graph that will display Lat vs Temp across a time period. Temp axis is only a point...will need more thought

**Resources:**

[**http://apps.who.int/gho/data/node.resources.api**](http://apps.who.int/gho/data/node.resources.api)

<http://www.denguevirusnet.com/epidemiology.html>

<https://en.wikipedia.org/wiki/Dengue_fever> (primarily in South America and Asia) would be good to display on a map...large geo.

http://worldmap.harvard.edu/data/geonode:country\_centroids\_az8

**Malaria**:

https://gateway.euro.who.int/en/hfa-explorer/

???

??

**Citations**?

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3066711/

**Assumptions:**

3 or more cases is considered significant-legitimate for a latitude

The focus of the exercise is to find the latitudes not the countries

The latitudes were calculated by….

Web links

|  |  |  |
| --- | --- | --- |
| Wilki pages for scraping |  |  |
| <https://en.wikipedia.org/wiki/Tropical_disease> |  |  |
|  |  |  |
| Web pages |  |  |
| <https://www.who.int/topics/tropical_diseases/en/> |  |  |
| <https://www.who.int/topics/tropical_diseases/factsheets/en/> |  |  |
| <https://www.niaid.nih.gov/research/neglected-tropical-diseases-types> |  |  |
| <https://www.cdc.gov/globalhealth/ntd/diseases/index.html> |  |  |
| <https://academic.oup.com/cid/article/66/6/815/4773343> |  |  |
| <https://www.kaggle.com/cdc/zika-virus-epidemic> |  |  |
| WHO API |  |  |
| <http://apps.who.int/gho/data/node.resources.api> |  |  |
|  |  |  |
| climate change data |  |  |
| temps | <https://www.ncdc.noaa.gov/cag/global/time-series> |  |
|  |  |  |
| specific disease geographic data | Example api call:  [http://apps.who.int/gho/athena/data/GHO/WHS3\_50?profile=xtab&format=html&x-topaxis=GHO;YEAR&x-sideaxis=COUNTRY&x-title=table&filter=COUNTRY:\*;REGION:\*](http://apps.who.int/gho/athena/data/GHO/WHS3_50?profile=xtab&format=html&x-topaxis=GHO;YEAR&x-sideaxis=COUNTRY&x-title=table&filter=COUNTRY:*;REGION:*) | <http://apps.who.int/gho/data/node.home> |
| yellow fever | <http://apps.who.int/gho/data/view.main.1520_50?lang=en> |  |
| Euro Health Org |  |  |
| Dengue fever | http://data.euro.who.int/cisid/?TabID=483584 | SAMPLE table: http://data.euro.who.int/cisid/?TabID=483585 |
| API calls | http://apps.who.int/gho/athena/api/COUNTRY?format=json |  |

Lesson learned

[{ year:int

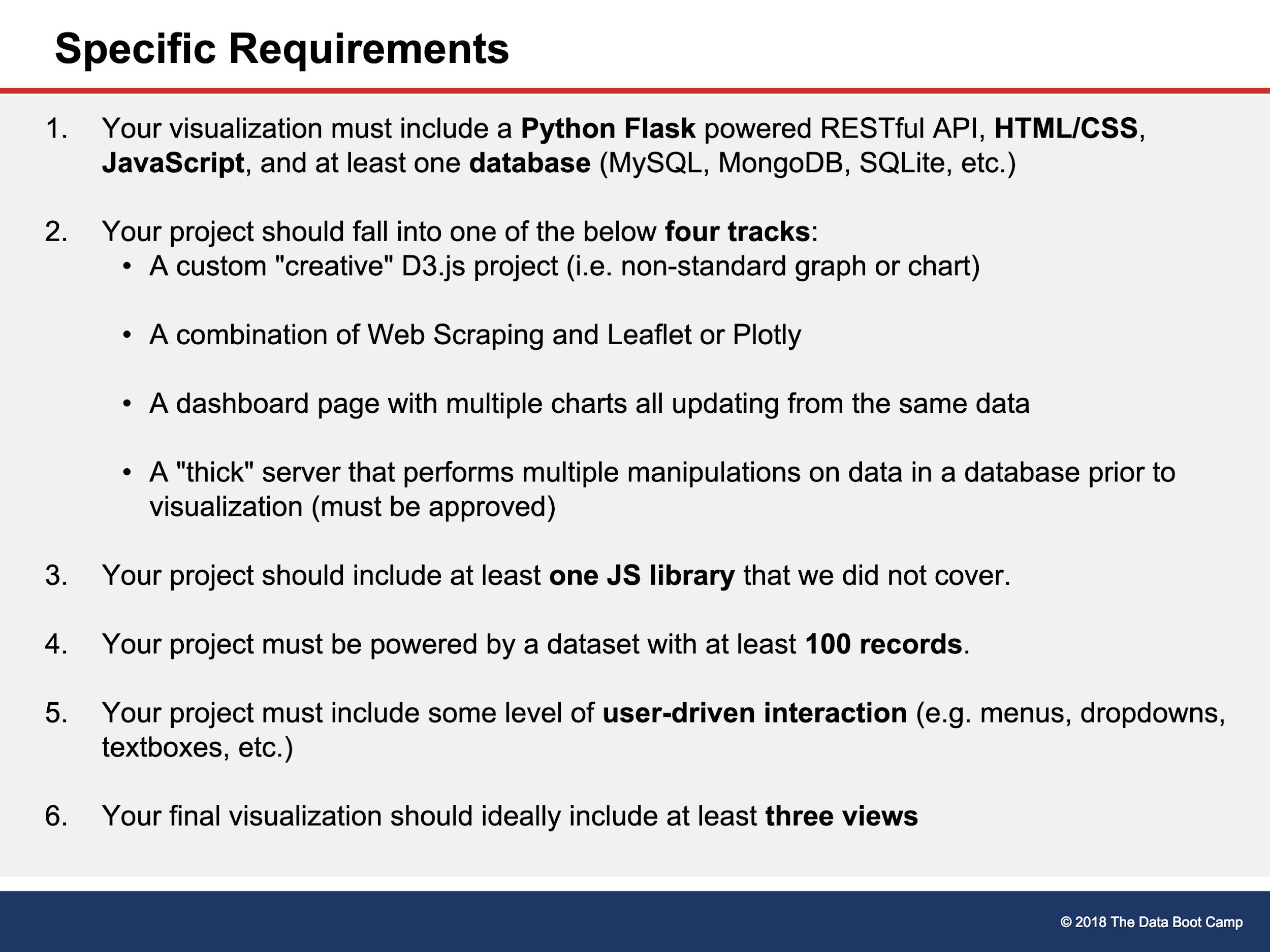
Temp:float

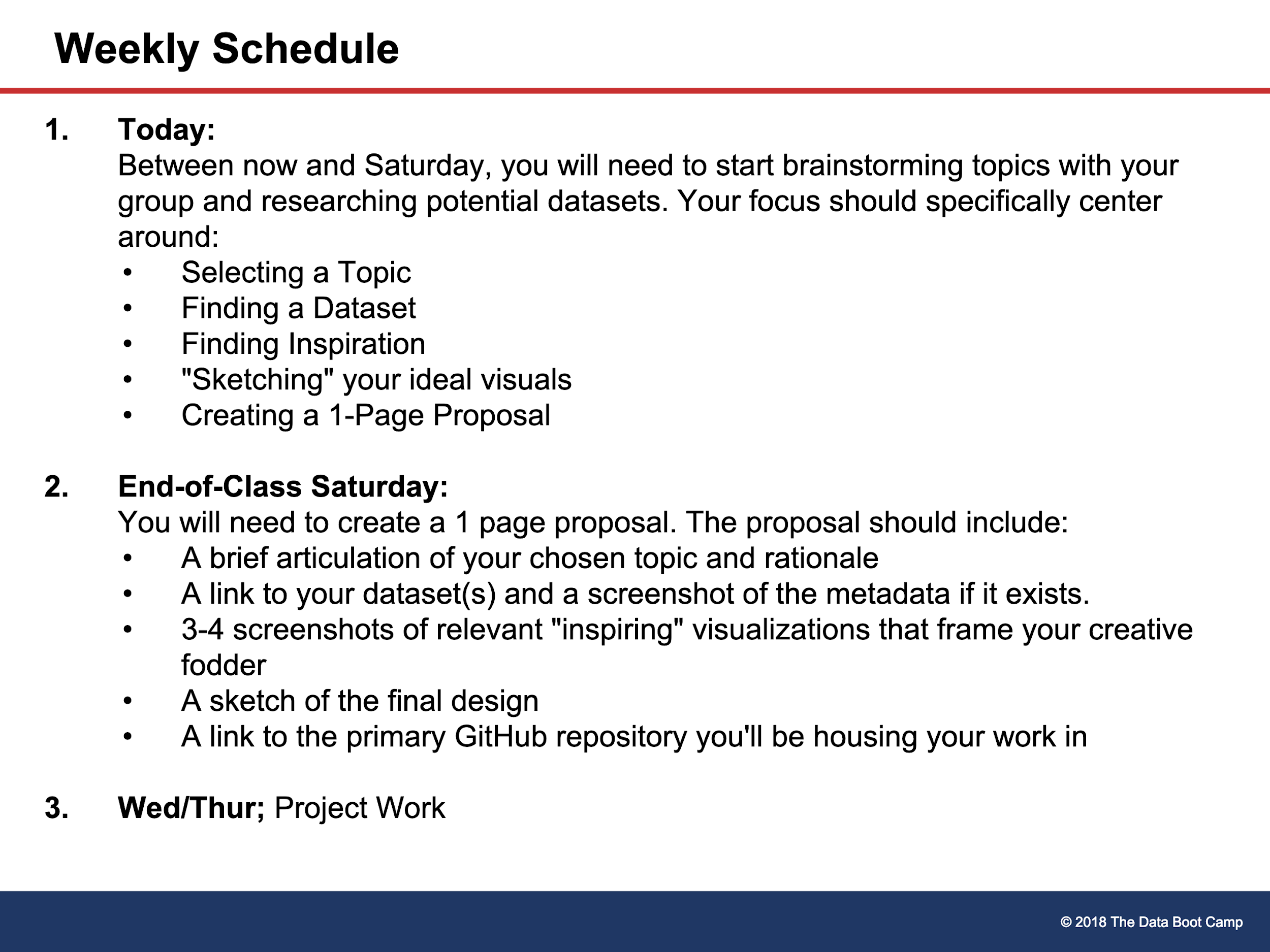
Diseases: [ {name string

Max lat float

Min Lat float}

]





Tasks for the next levels:

**Draft Proposal:**

‘Iadhfgoiadhfgo’iadnhfnhgo’

NOA API

https://www.ncdc.noaa.gov/cdo-web/api/v2/{endpoint}