

## **Background and context**

We are planning to build a country metrics (GDP, population, etc) predictor. We plan on building a machine learning algorithm that can generate future metrics using training data from 1956 - 2017. Our chosen metrics will be weighted into one easy to understand coefficient which we will attempt to forecast. We also plan to display this data for all countries in a GUI (maybe website?) that is also interactive. Users will have the ability to manipulate the weights of the metrics they value more (ex. high population is important, thus more weight). The GUI will display the according coefficient depending on the inputted weights of the user.

The program will have two components the machine learning aspect and the GUI/website aspect.

The machine learning model will be responsible for using prior data to generate future data. The GUI will be responsible for displaying the outputs. The interactive user interface make changes to the model, which will then change the displayed metrics.

### **Machine Learning Progress:**

- Able to read a CSV file, extract data, and create a dictionary with country names as keys and a list of a given statistic (GDP, population, etc) starting with 1960 as the first year.
- Able to create a simple linear regression with a given list of data. Our code currently can not handle missing data, which is a problem we are trying to solve.

### **GUI Progress:**

- Have developed a simple website using HTML, CSS, and Javascript.
- Has workable sliders.
- Can display data from a Python list.
- Has dropdowns!

## **Key questions**

We want to ask questions regarding two areas of discussion:

### **Machine Learning:**

1. We are planning to use a basic linear regression to predict future values, what is the best and easiest way you know to do this?
2. We are having a problem with None values, do you know a method to build a linear regression with null values?
3. What indicators are you interested in seeing for the “coefficient”(the weighing value for the map colors)?

GUI:

1. Google Visualization can only take in 2 data points to display on the map (i.e. coefficient and population) Are there other libraries you know of that better handle larger quantities of data?
2. How do you feel about the general interactivity of the website, currently? How can we better lay out the website?
3. What type of information would be useful to the website user to know, and how best can we share that info? (I.e. blurb on this class? Background of the idea?)

Agenda for technical review session

0-2 Project Overview

2-5 Current Progress

5-15 Group Breakout and Q&A

Machine Learning:

5-8 Question 1

8-11 Question 2

11-15 Question 3

GUI:

5-8 Question 1

8-11 Question 2

11-15 Question 3

Survey: <https://tinyurl.com/SampeAndSara>

Presentation:

[https://docs.google.com/presentation/d/1AcoKq0RrW\\_PZlZ8ZJw3gA60mzpdYgft-4dynlAPNnY/edit?usp=sharing](https://docs.google.com/presentation/d/1AcoKq0RrW_PZlZ8ZJw3gA60mzpdYgft-4dynlAPNnY/edit?usp=sharing)