John Traylor

SI 507

3/25/2018

Final Project

**Proposal**

The Peace Corps is an American government agency that sends volunteers to developing countries to foster relationships with foreign governments and people through international development projects in several work sectors: Public Health, Education, Water & Sanitation, Agriculture, etc.

This project will collect data on the countries in which Peace Corps volunteers serve and the languages that they learn in order to integrate into their host communities.

**Data Sources:**

1.) *Scrape* data from the Peace Corps Countries webpage (4 Challenge points)

\*Webpage: <https://www.peacecorps.gov/countries/>

* Country Name
* Country Region
* Number of current volunteers
* Work sectors in which volunteers operate
* Link to country specific webpage

2.) *Crawl* to country specific webpage to *Scrape* additional data (4 Challenge points)

\*Example Webpage (There is a webpage for each country in which Peace Corps operates): <https://www.peacecorps.gov/benin/>

* Volunteers to date
* Official Language

3.) *Json Parsing* the “countries.json” file from Project 3 (1 Challenge point)

* Country Latitude
* Country Longitude
* Population

**Total Challenge Points: 9**

This data will be used to create the following two tables:

**Countries:**

Country Name: TEXT, Primary Key

Country Region: TEXT

Number of Volunteers: INTEGER

Official Languages: TEXT, Foreign Key

Work Sectors: TEXT

Peace Corps Country URL: TEXT

Volunteers to Date: INTEGER

Country Latitude: INTEGER

Country Longitude: INTEGER

Country Population: INTEGER

**Languages:**

Language: TEXT, Primary Key

Countries where spoken: Text, Foreign Key

Number of Current volunteers speaking: INTEGER

Number of Past Volunteers speaking: INTEGER

**Data Presentation:**

I will use Plotly to present my data in the following four ways:

1. World Map Scatter Plot that displays a point for each Peace Corps country with a hover caption that displays number of current volunteers and the languages they speak.
2. Bar graph that displays the top five countries by volunteer / country population ratio.
3. Pie chart that displays languages spoken by volunteers by percentage
4. Bar graph that displays the volunteer population of world regions.

\*\*Instead of presentation method #4, could I allow users to launch a web browser that would look at a country of choice’s specific Peace Corps webpage? For example, if a user would like to see more about Peace Corps Benin, they could go to

< <https://www.peacecorps.gov/benin/> > via my interface?

Could this count as one of my four data presentation methods?