

---

# Capstone Project – Environmental Standard and Its Correlation on National Employment, Wealth Inequality and Happiness Index

**Presented by Andy Nguyen**

BrainStation Data Science program

June 7, 2020

---

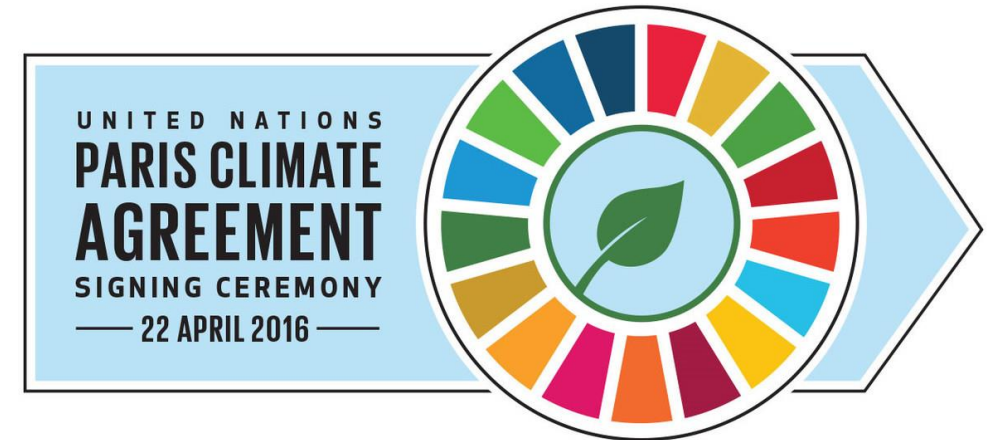
# Introduction

## Key Questions and Metrics

### Global Environmental Crisis Key Metrics<sup>(1)</sup>

Percentage of reduction of the world real GDP per capita by 2100, without major breakthrough	7.22%
The second-hottest year on record	2019
Average sea level rise since 1880	23 cm

### Key Global Environmental Progress



### The Debate of Environmental Impact on The Economy



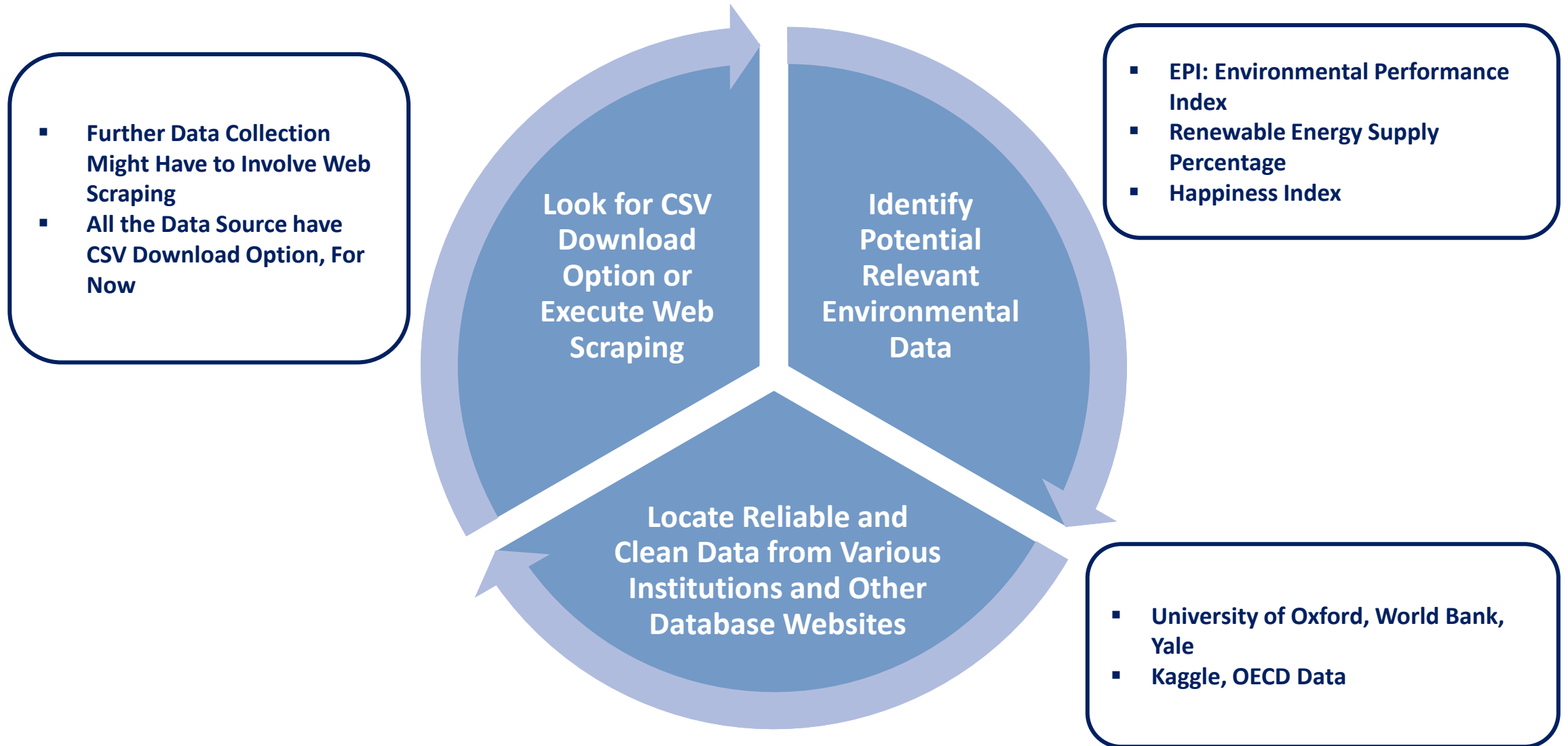
The US withdraws itself from the Paris Agreement in 2017 led by Donald Trump, ceasing all participation in the global effort to slow global warming

The debate evolves around whether taking strong actions against climate change would have an expensive cost, specifically on government spending, job creation and social security

(1) Source: US Government Database

# Data Collection

## Source and Process



# Data Description

## Column Presentation and Introduction

Code	Country	Year	Happiness Score	Unemployment Rate %	EPI Score	Annual CO2 emissions (tonnes)	Renewable Supply %	Fossil-fuel pre-tax subsidies (\$USD)	RE Consumption (% of Total)
AGO	Angola	2017	3.79	7.11	29.7	3.54E+07	47.25	9.64E+08	49.56

### Target: Happiness Score

- This column represents the happiness index of 119 countries around the world, calculated by respondents of each nation's evaluation of their living quality from 0 to 10

### Target: Unemployment Rate %

- This column represents the unemployment rate as the percentage of the labour force of each country

### Feature: EPI Score

- Environmental Performance Index (EPI) ranks countries on environmental health and ecosystem vitality.
- EPI offers a scorecard that highlights leaders and laggards in environmental performance and provides practical guidance

### Feature: Annual CO2 emissions tonnes

- This column measures the CO2 emission rate of each nation into the atmosphere

### Feature: Renewable Supply Percentage

- This column represents the percentage of energy supply of each nation that comes from renewable energy

### Feature: Fossil-fuel pre-tax subsidies (\$USD)

- This column represents the amount of government subsidies that flow towards fossil fuel industry

### Renewable Energy Consumption (% of Total)

- This column represents the consumption of renewable energy as a percentage of total energy consumption in each country

# EDA and Modelling

## Process and Preliminary Results

### Exploratory Data Analysis Process

#### Step 1

- Load all the CSV files into Jupyter Notebook and examine the columns and rows
- Rename some of the columns (Country, Code, Year) of all data sets to keep things consistent

#### Step 2

- Select rows that are in 2017 to keep the data consistent for now, since some of the data only have value up until 2017
- Remove NaaN rows

#### Step 3

- Merge all the Data Frame together using the column "Code" as the key
- Remove non-linear columns with the target and high VIF ones.

There are 119 rows and 10 columns in the final DataFrame

### Modelling Preliminary Result and Analysis

#### OLS Regression Results

Dep. Variable:	Happiness Score	R-squared	0.643
Model:	OLS	Adj. R-squared:	0.634
	Coef	Std Err	P> t
const	3.6905	0.259	0
Unemployment Rate %	-0.0588	0.011	0
EPI Score	0.0473	0.004	0
Renewable Supply %	-0.0024	0.002	0.31
Sklearn Training Score		0.71	
Sklearn Test Score		0.43	

#### OLS Regression Results

Dep. Variable:	Unemployment Rate %	R-squared	0.02
Model:	OLS	Adj. R-squared:	-0.01
	coef	Std Err	P> t
const	9.0442	2.138	0
Annual CO2 emissions	-4.93E-10	5.14E-10	0.339
Fossil-fuel subsidies	-2.06E-13	6.09E-11	0.997
EPI Score	-0.0178	0.034	0.598
Renewable Supply %	-0.027	0.021	0.198

There is no correlation between job creation and environmental standard

Adj R<sup>2</sup>: -0.01

# Plan of Action

Stay Tuned

