**Lesson 3 – Little Data to Big Data**

**\*\*Instructions:** Please change the text color of your responses to red text. Please organize the endings to each page.

**ACTIVITY 3.2.2 – Shocking Data Trends**

Observe the data and identify the following information. What year range does this data contain?

|  |
| --- |
| The data contains the years 1990 to 2016. |

Approximately how many countries are represented in this data set?

|  |
| --- |
| The data set contains 257 unique items and 216 country codes, but some of these are territories of other countries, such as the US Virgin Islands. |

Is every country in the world represented in this data set?

|  |
| --- |
| No. |

Why do you think some countries might not be represented in this data set?

|  |
| --- |
| It may be impossible to gather data in some countries due to bad infrastructure or other conditions. |

Describe an algorithm you would follow to compare two different countries without programming.

|  |
| --- |
| I would graph the two countries’ graphs and see which one has a higher line. |

Break down and describe what is happening with the following code.

sum\_elec = df[df['Entity']==c]['Access']

|  |
| --- |
| df[‘Entity’]==c is a conditional that checks if the entity, or first column, for each row, is equal to c.  df[df[‘Entity’]==c] gets all row in df that match this requirement.  [‘Access’] gets the Access column from these rows, which is the fourth column. |

Identify the parts of the code that create each of the following:

1. The data set that will be used to generate the graph
2. The countries that will be plotted
3. The labels for the x-axis, y-axis, and the title

|  |  |
| --- | --- |
| 1. | df = pd.read\_csv("elec\_access\_data.csv", header=0) |
| 2. | my\_countries = ['United States', 'Zimbabwe','Cuba', 'Caribbean small states', "Cameroon", "Burundi"] |
| 3. | plt.ylabel('Percentage of Country Population')  plt.xlabel('Year')  plt.title('Percent of Population with Access to Electricity') |

**Conclusion**

Consider the digital divide you learned about in lesson 2.2. How does access to electricity affect a country's access to computing innovations?

|  |
| --- |
| Electricity is universally required to use computing innovations, as most, if not all, modern computing innovations rely on the use of electrically-powered circuits. |

How can the output of data be used to affect global change?

|  |
| --- |
| Spreading data can drive people to help global change, as making people more aware of situations makes it more likely for them to want to help solve problems. |