NOVA IMS Information Management School

MAA

Master Degree Program in Data Science and Advanced Analytics

DATA VISUALIZATION REPORT

Visualizing the CO2 Emissions Among Different Continents in the World

Group Members:

Foazul Islam-M20200750
Angira Halder-M20210436
David Santos-R20181082

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

1. INTRODUCTION

Since the Kyoto Protocol, signed in 1997 and active from 2005 to 2012, there have been many agreements in every scale (national, continental and worldwide) with the objective of resolving the climate emergency, much of them (especially in the beginning) focusing on reducing greenhouse gas emissions that go into the atmosphere and disrupt the ozone layer, thus being the main cause for climate changes. With that in mind, our aim is to understand how the different continents have been working towards those goals by looking at data for one of the greenhouse gases, CO2, emissions from 2011 and 2020. We will explore the variation in those values between these years, as well as the average emissions and cumulative ones, through visualization techniques provided by Python and its Plotly library.

2. DATASET DESCRIPTION

To help us analyze our problem, we used a simple dataset, containing data on CO2 Emissions by each continent (Asia, Africa, Europe, North America, South America and Oceania) between the years of 2011 and 2020, as seen in **Table 1.** Also, in order to perform an analysis on the variation of CO2 emissions between the selected years by the users, the same dataset was pivoted, thus showing the continents as the index values, and the values per year in the following columns, as seen in **Table 2**. Due to an error in the dataset, a change had to be made through code. Both datasets are available in the GitHub link given at the end of this report.

Continent	**Year**	**CO2_Emission**
Asia	2011	17746631804
…	…	…
Asia	2020	20317058379
Africa	2011	1258708905
…	…	…
Africa	2020	1326043539
Europe	2011	6023668149
…	…	…
Europe	2020	4946034489
North America	2011	6771607912
…	…	…
North America	2020	5775158655
South America	2011	1103963494
…	…	…
South America	2020	994160327
Oceania	2011	451765877
…	…	…
Oceania	2020	444397224

Table 1- Main Dataset

Continent	2011	…	2020
Asia	17746631804	…	20317058379
Africa	1258708905	…	1326043539
Europe	6023668149	…	4946034489
North America	6771607912	…	5775158655
South America	1103963494	…	994160327
Oceania	451765877	…	444397224

Table 2-Pivoted Dataset

3. DATA VISUALIZATION

For visualizing the data, an HTML dashboard is created, and it can be divided in two parts: the interactive section, where the user can filter some information like the continent or a time range, using some Dash Core Components, from the Plotly library, and three plots illustrating the information filtered by the user.

3.1. Interaction

For the interaction section, three components where created: a dropdown menu to select the continent for one of the graphics (with the possibility to select multiple options, in order to allow a comparison between continents), a range slider to select the range of years to analyze in two other graphics (with a minimum one year distance between values, guaranteeing that two years are selected and not only one) and two radio buttons, for the user to choose weather they want to see one of the figures show the mean emissions from the continents, or the variation between two years (selected in the range slider). This section is illustrated below, in **Figure 1**.

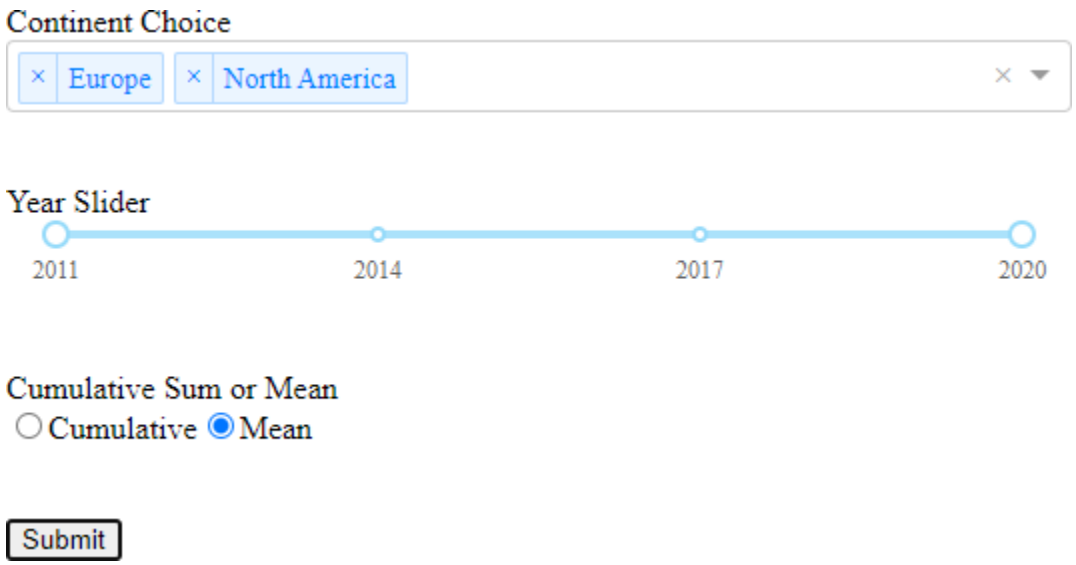

Figure 1- Interactive Section of the Dashboard

3.2. Plots

Three plots were created to help the user understand the data: one simple scatter plot showing the emissions of each selected continent (from the dropdown) along the years, a bar plot that gives the user the chance to see the mean emissions from all the continents between the selected years (from the range slider) or the cumulative emissions of every continent between the same years, and another bar plot showing the variation between the emissions of the two selected years in the interaction section. All the plots can be seen below in **Figures 2, 3 and 4**, respectively.

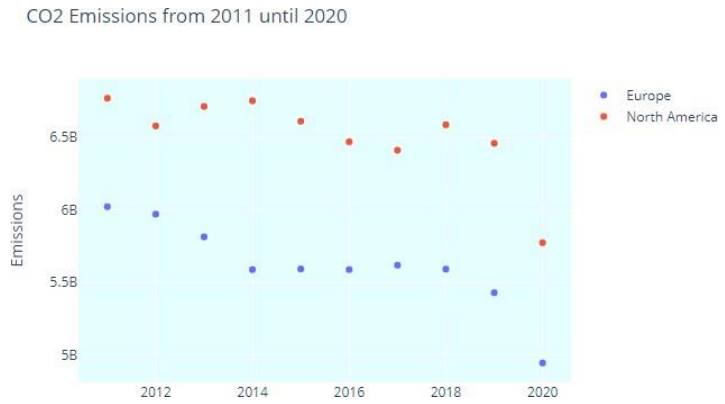

Figure 2- Scatter Plot

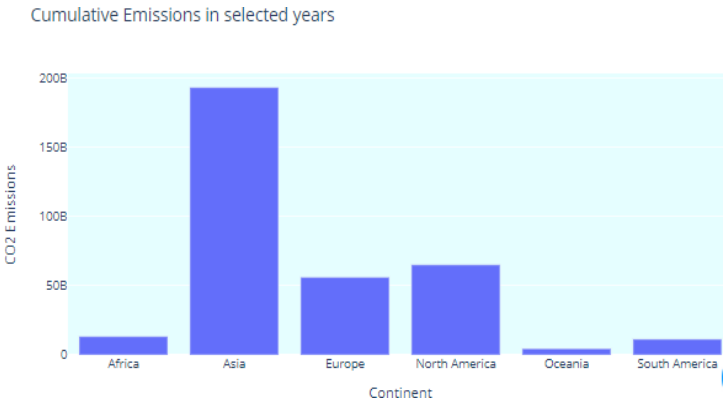

Figure 3- Bar Plot for Cumulative Emissions

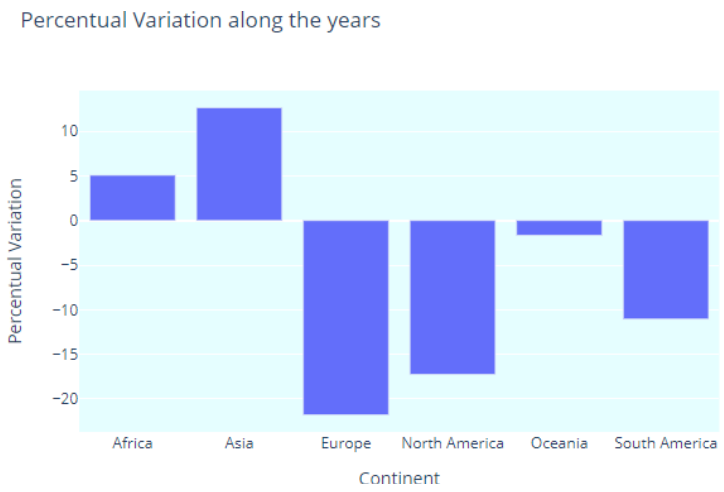

Figure 4- Bar Plot for Emissions Variation

3.3. Data Encoding

As aforementioned, the dashboard was created with several features, of both visualization and interaction and filtering. These features came from Dash, a low-code framework written on top of Plotly to create data apps. Some of these components include the Dropdown, used to select the continents to visualize, the Graph, used to build the main visualizations, in this case, a scatter plot and two bar plots, the Radio Items, used to select the type of calculation to perform in one of the visualizations (cumulative sum of mean), and the Range Slider, used to select the range of years in the cumulative sum or mean, or the two years to calculate the variation from.

4. CONCLUSION

After applying different types of visualization techniques dealing with our dataset, we can realize the progress regarding CO2 emissions of each continent (except Antarctica) and compare these progresses in an absolute and relative ways, having in mind that the continents have different sizes, and thus emit different levels of CO2.

We can see that some continents like Europe, North America and South America have been working hard in reducing these emissions, showing a decreasing tendency in most of the years. It should also be mentioned that Europe and North America had very similar values for every year, being the two most similar continents. Then the remaining continents, Asia, Africa and Oceania, show a positive tendency, meaning that during most years, they have increased their emissions. Although, it should be noticed that every continent showed a decrease in emissions between 2019 and 2020, the last years recorded in our dataset, giving some hope to this international effort of guaranteeing a better tomorrow for future generations.

For future work, we should consider using a more complex dataset, given that the simplicity of this one was the biggest limitation to our work, with data from countries rather than continents, information regarding other greenhouse gases emissions besides carbon dioxide, and a wider range of time to analyze, in order to understand where these numbers are coming from, and if the tendencies observed were continued in 2021 and 2022.

References:

Evergreen, S., & Metzner, C. (2013). Design principles for data visualization in evaluation. *New directions for evaluation*, *2013*(140), 5-20.

Source of Datasets: https://ourworldindata.org/

Dash Documentation: https://dash.plotly.com/

Dash Running app on:

https://dv-final-project.onrender.com/

GitHub Link: https://github.com/foazul/DV_FINAL_PROJECT