Introduction

Climate change has been a topic of discussion for decades. With the rise of natural phenomena such as unusual heat waves, floods, and droughts, people have become more aware of their impact on the climate. Air emissions significantly contribute to human-induced climate change driven by modern-day consumption patterns.

Air emissions refer to the gases and particles released into the atmosphere from various sources. Among these, the increase in greenhouse gases such as carbon dioxide and methane pose a significant risk to the environment and is a major cause of climate change. In addition to greenhouse gases, airborne particles also threaten human health. Analyzing targeted data can yield valuable insights that help mitigate the effects of climate change.

The Federal Statistical Office of Germany provides valuable data detailing the emissions from various economic sectors and their evolution over the years 1995-2021.

Question

Which industries emit the most air pollutants and greenhouse gases over the years in Germany?

The focus of this project is to understand the amount of air emissions produced by each economic sector in Germany and identify the sectors with the highest emissions over the years. Identifying the economic sectors that generate the most air emissions, we can better target reduction efforts and raise awareness about our consumption habits.

Data

Licensing

Data is provided by The Federal Statistical Office of Germany (Statistisches Bundesamt). According to information in here, after 2020, the source of publication has changed to GENESIS-Online database, which can be accessed in here. According to Genesis-Online, data is licensed under the "Data Licence Germany - Namensnennung Version 2.0", licence text available at www.govdata.de/dl-de/by-2-0. Data can be used, altered, processed, and merged as long as the user ensures the name of the provider, the link to the dataset, and refers to the license text, which are fulfilled in the project-plan.md file, scripts, and the report.

Structure and the Meaning of Data

Data is presented in tabular form. Air emissions between 1995-2019 can be accessed from one source, whereas 2020 and 2021 are accessed each from different sources. Data includes year information, types of gases, and different economic sectors. With this information, one can read, for example, the amount of carbon dioxide, emitted by the energy supply sector, in year 2005. The result for this query would be almost 393 million tonnes.

An important thing to mention is how to interpret the big numbers represented in the data. The documentation explains that the data uses the metric “carbon dioxide equivalent” (CO2-eq), often expressed as million metric tonnes of CO2-eq. CO2-eq is used to compare the emissions from various greenhouse gases based on their global-warming potential (GWP), by converting amounts of other gases to the equivalent amount of carbon dioxide with the same global warming potential. For example, the documentation states methane has 25 times the impact on climate, in comparison with CO2. That means, 1 million metric tonnes of methane is equivalent to 25 million metric tonnes of carbon dioxide. Also, the

data is reported in units of 1000 tonnes, meaning the result of the above query is actually 393 billion tonnes of CO_2.

Analysis

Methodology

The posed question of the project suggests our interest lies in proportions rather than the absolute amount of emitted gases. The unit used in the data (GWP) allows us to consolidate information into a single measure called total emissions. Considering these aspects, in each year's data for every economic sector, the emissions of gases are aggregated to create a new metric termed "Total Emissions". Certain gases were excluded from this calculation, mainly due to the introduction of new gases in the 2020 and 2021 datasets and the absence of some gases from previous years. However, these discrepancies have negligible effects on the overall outcomes since the non-matching gases contribute insignificantly to total emissions (in most cases less than 0.01%) and do not influence the ranking we seek. However, the other gases are not disregarded; they are still included to the equation when calculating proportions.

Given that there are nearly 50 economic sectors in the dataset, including all of them in a single plot is impractical. Therefore, only the top 7 sectors are displayed in the graph, while the remaining sectors are aggregated into an "Others" category.

Furthermore, it's important to note that the data for 2020 and 2021 employ distinct naming conventions and economic sector definitions. According to the documentation, emissions from power plants are now allocated to their respective economic sectors, whereas previously they were allocated to the energy supply category. Consequently, while emissions from 'electricity supply' have decreased, as an example, those from 'chemicals' have increased when the results over the years are plotted. Nevertheless, these discrepancies do not significantly alter the ranking, allowing us to interpret the results effectively. This change is illustrated in Figure 2.

The data processing steps mentioned above results in two main features: total emissions and proportions. Total emissions are used to visualize evolution of a sector's emissions over the years, while proportions help identify the sectors making the greatest contributions.

To determine which sectors contribute most to air emissions, data is sorted in descending order based on their total emissions for each year. The rank of each sector is then summed across all years. For example, if "private households" ranks first in one year and second in another, the sum would be 0 + 1 = 1 (considering zero-based indexing). The sector with the lowest average ranking indicates the highest emissions. This method identifies sectors that consistently emit high levels of emissions over multiple years, rather than focusing on single-year outliers. By summing the ranks across years, we capture a long-term trend which is more indicative of a sector's impact on air quality.

Results

When years are compared with each other, we can see that they have similar distributions. For example, figure 1 illustrates the proportions in 2008 and 2018.

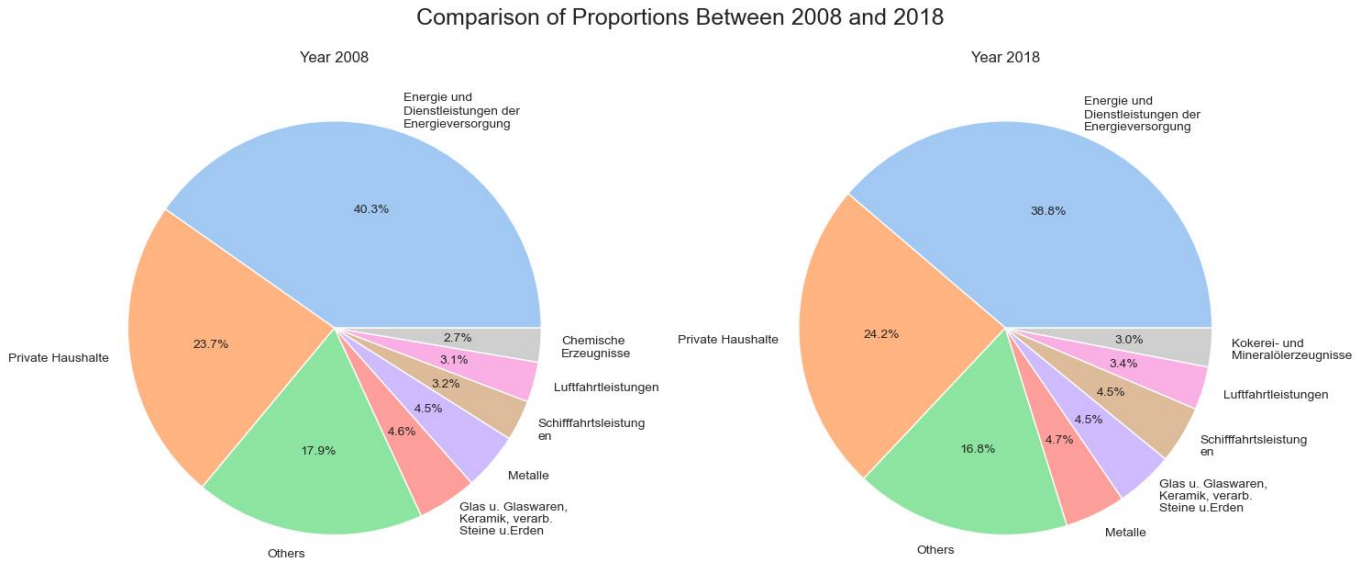

Figure 1: Comparison of Proportions Between 2008 and 2018

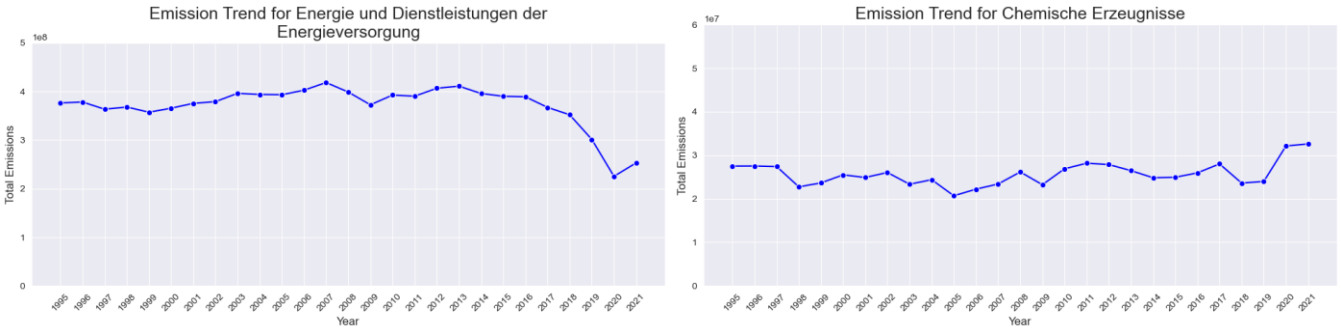

Figure 2: Illustrating the effect of different measuring methods

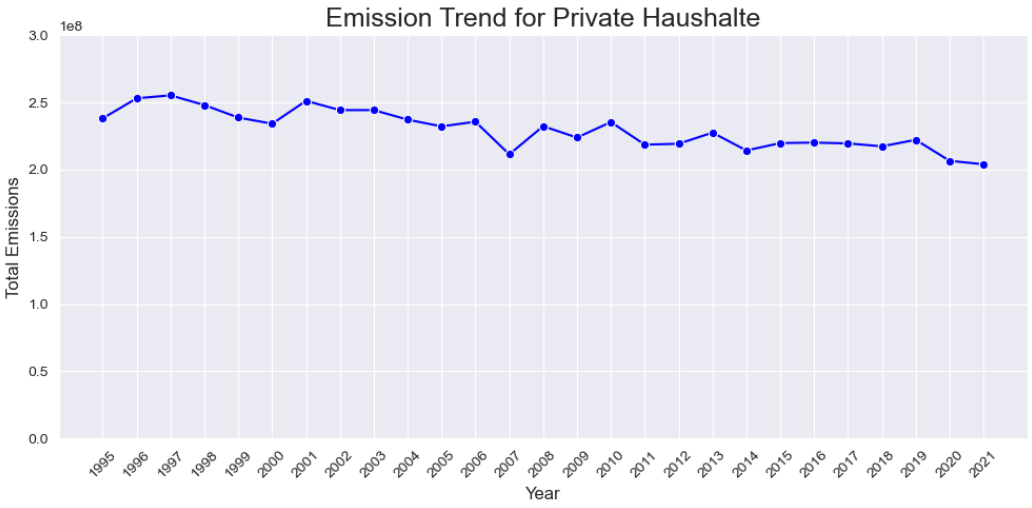

Figure 3: Emission trend for Private Households

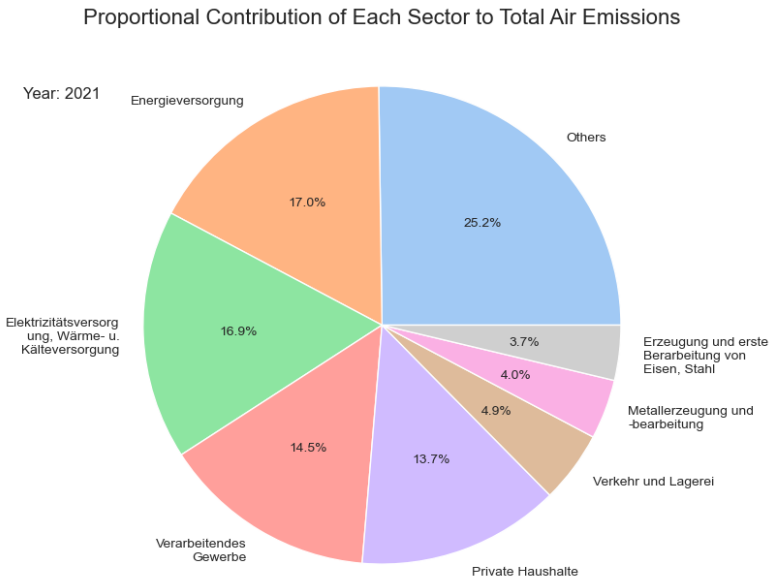

Figure 4: Proportional Contribution of Each Sector to Total Air Emissions in 2021

In figure 2 and figure 3, the trend of air emissions over the years, or in other words how the amounts have changed annually is illustrated.

In figure 4, we can observe the new sector names that are introduced by the update in 2020 and their distribution.

Interpretation

Based on the figures above, it is evident that the proportions of economic sectors have remained largely consistent over the years, with no significant fluctuations in emission trends. The top 4 to 5 emitters consistently maintain their positions, though rankings among other sectors can vary. For example, chemical-related emissions ranked among the top in 2008 but did not maintain this position in 2018, as shown in Figure 1.

In figure 4, the proportion of “private households” is notably smaller compared to previous years. This is attributed to the introduction of 17 new economic sectors in 2020 and 2021. However, as shown in figure 3, there was no significant decrease during these years.

It is observed that energy supply, private households, manufacture of non-metallic mineral products, metal related emissions, air and water transport, and manufacture of petroleum products have the biggest proportions among all the sectors.

Conclusion

When the method mentioned in the section “methodology” is used, the ranking below is obtained:

1. Electricity supply / power supply
2. Private households
3. Metal related emissions
4. Manufacture of non-metallic mineral products
5. Air transport
6. Chemical production
7. Manufacture of coke and refined petroleum products

The findings indicate that the primary sources of air emissions in Germany are electricity production, private household consumption, and metal production and processing. These results align with expectations, as stated here by the Statistical Office of Germany, fossil fuels remain the dominant energy source for electricity generation in the country.

Limitations

The data allows us to analyze trends from 1995 to 2021. Notably, we lack information on emissions from industries before 1995, which can be significant given that climate change and air emissions were less prioritized, and regulations were less stringent at that time. Additionally, we cannot assess recent changes in air emissions. Another limitation is our inability to conduct detailed analyses at city or monthly levels, which could hinder targeted efforts and timely interventions in cities with higher air pollution levels.