# Visualizing Pollution Data on European Industries

Visual Analytics

(CDSCO1003U)

Ivana Jasna Caltagirone, Tommaso Ghisini,

Frederik Gøtske, Arianna Sammarchi

Copenhagen Business School

06 January 2023

24 pages, 31,077 characters

# Contents

# 1 Introduction

There is a growing demand for industries to measure and understand their environmental impact, as concerns about climate change and corporate responsibility continue to rise. However, collecting and analyzing the vast amount of data that is available on this topic can be a complex and time-consuming task, and a dashboard that is designed to summarize and visualize this data on a large scale can be a valuable tool for understanding and addressing corporate environmental impact. In this project, the focus is narrowed to insights at the European level, due to the availability of data and the scope of the research, as well as to make it more relevant and useful for the target audience, the Environmental Council of the European Union. The council plays a key role in shaping EU environment policy, including efforts to use resources responsibly and safeguard human health. It also addresses international environmental issues, particularly those related to climate change (European Council, 2022).

The aim of this project is to create a dashboard that provides a comprehensive overview of the environmental impact at the European level, allowing policymakers to identify and target areas and industries with significant impact in order to develop new policies and solutions that enhance sustainability. This tool will allow users to compare and visualize the environmental impact of multiple industries and countries, as well as break it down into the particular affects of each company in each year.

The Dashboard can be accessed at the following link:
**Link to the Tableau Public Dashboard** (A.1)

# 2 Approach

The main data source for the dashboard is *Corporate Environmental Impact: Measurement, Data, and Information*, a study conducted by David Freiberg et al. (2021). The study provides a novel and thorough method for assessing and monetizing corporate environmental impact, making it a potentially powerful tool if presented in the right way for the right people. The study features a comprehensive data set comprising environmental impact metrics, calculated using emissions data, for global enterprises. These metrics encompass a range of environmental

impact areas:

- Human Health (*Working Capacity*)
- Abiotic Resources
- Biodiversity
- Crop Production Capacity
- Fish Production Capacity
- Meat Production Capacity
- Water Production Capacity
- Wood Production Capacity

The authors sum all areas to create the Total Environmental Cost (TEC) for every entry. This measure refers to the costs associated with the negative impacts or externalities of a product or activity on the environment. These costs are often not reflected in the price of a product or service, but they can still have significant economic and social impacts. The main reasoning behind the Freiberg et al. (2021) study is to create a measure that can be used to fill this gap.

The measures are heavily based on corporate emissions data, combined with the Environmental Policy Support (EPS) database and the AWARE model. EPS provides a method for estimating the monetary value of an organization's environmental impacts, such as emissions, through monetization. This involves using monetization factors, which are based on the concept of willingness-to-pay, or how much people are willing to pay to avoid or mitigate negative environmental impacts. The EPS database does include global water monetization factors, but the AWARE model can provide additional factors that consider local water scarcity. These factors allow for more accurate comparisons of water use across countries with different levels of water availability and enable the use of a global price once local water use is converted to an equivalent global value (Freiberg et al., 2021).

To account for differences in firm size, Freiberg et al. (2021) uses revenue as a proxy to scale TEC, thereby creating the Total Environmental Intensity measure. As the dashboard can be instructed to display data for any of the environmental domains, it uses the generic terms *Cost* and *Intensity* to denote the Cost and Intensity for any selected impact area. This paper will use the denotions Cost of Environmental Impact (CEI) and Intensity of Environmental Impact (IEI) to indicate the cost and intensity of any selected impact area.

In addition to the main data, the Environmental Risk Score, which tries to indicate the impact of a company's operations on the natural environment similarly to the main data set, from

Sustainalytics ESG Data provides policy makers with an alternative measure of environmental risks. As ESG has gained widespread use in corporate reporting and sustainability analysis, it can be utilized for further exploration of the issue at hand and as a sanity check for the main data, especially as the measures used to calculate it differ significantly from those of the main data set (et al., 2021). Unfortunately, the scope of the task does not allow the collection of ESG data for all years, and as such, only data from 2019 is included. While still a valuable measure, including only one data point per year limits the ability of users to identify any trend in the data. Furthermore, there is only ESG data for 195 out of 757 firms due to data availability.

The tool stack used to complete the project is composed of Python, Jupyter Notebook and Tableau. Python not only allows for the preparation and cleaning of data in an efficient way, but also to scrape the additional data needed for the dashboard. Furthermore, Jupyter notebooks provide a powerful platform for writing and running Python code, as well as for documenting and sharing the work across the team. Tableau is a good choice for creating visualizations as it is easy to use, highly customizable, and can create charts, graphs, and dashboards that are visually appealing and informative. It integrates well with a variety of data sources and has advanced features for data analysis.

# 3 Target Audience

By centering the analysis around European data, the primary focus of the dashboard is policy makers operating at the management level in environmental assessment, such as the Environmental Council shaping the Environmental Policies' frameworks for the European Parliament (European Parliament, n.d.).

The inclusion of time-based data, as well as the ability to select specific regions, countries, and industries for analysis and comparison, makes the dashboard a valuable tool for assessing the impact of regional and industrial policies as well as to identify areas that should be prioritized in future legislation. Furthermore, corporate environmental impact is of interest to a variety of stakeholders such as the European Environment Agency which provides independent information for the policy-makers at the European Level and the general public (European Environment Agency, n.d.).

A very different type of user is the institutional investor, especially those focusing on sustainability or hedging against environmental risks, such as future legislation. As an organization's environmental impact can affect financial performance, competitiveness, and reputation, the dashboard allows investors to measure and manage their own environmental impact, make informed investment decisions, and benchmark the performance of different companies and industries.

# 4 Data Preparation

A detailed overview of the data preparation process, including the various steps taken to clean and organize the data is provided in the following section.

## 4.1 ETL Process

The main steps undertaken in the ETL process are:

- Loading, Exploration and Filtering
- Feature Engineering: macro industries and revenues
- ESG Data Extraction
- Final Cleansing and Export

### 4.1.1 Loading, Exploration and Filtering

The initial data set consists of 14035 entries with 34 attributes. The attributes can be grouped as presented in Table 1.

Several of these attributes (*SDGs, GICS Subindustry, EI (Operating Income), Imputed %*) are outside the scope of our project and dropped from the data set. Additionally, 21 entries are missing from the *Country* attribute and are thus removed.

In order to focus on the European continent, all non-European entries are excluded. It is worth noting that Russia, which is partly located on the European continent, is excluded from the data set for several reasons. First, Russia has a relatively low number of firms compared to other European countries. Second, Russia has a complex and often contentious relationship

| | **Attributes** |
|---|---|
| **Instance Information** | Year, Company Name, Country |
| **Industry Information** | Industry (Exiobase), GICS Subindustry |
| **Environmental Intensity** | Environmental Intensity (Sales), Environmental Intensity (Operating Income) |
| **Environmental Cost** | Total Environmental Cost |
| **Environmental impact areas** | Abiotic Resources, Biodiversity, Crop Production Capacity, Fish Production Capacity, Meat Production Capacity, Water Production Capacity, Wood Production Capacity, Working Capacity |
| **SDGs** | 16 attributes not yet thoroughly described by the authors |
| **Imputed %** | The percentage of raw data used to calculate the various environmental impacts from a given entry that was imputed by the authors |

Table 1: Attributes of raw data set (et al., 2021)

with other European countries, which can create uncertainty about the quality of data available. Finally, Russia is not politically influenced by the European Union in terms of environmental policies. While the EU has implemented strict environmental regulations that impact businesses operating within its borders, Russia has its own regulatory framework. Additionally, companies based in the Isle of Man, Jersey, and Gibraltar were reclassified to the United Kingdom due to their small sizes and deep connection. The initial data filtering process reduced the sample size to 4532 entries with 14 attributes.

### 4.1.2 Feature Engineering: Macro Industries and Revenues

Feature engineering is mainly performed around two dimensions: industries and revenues. The companies in the data set are classified into 51 distinct industries following the Exiobase Industry classification (Consortium, n.d.), which provides a detailed and exhaustive framework. However, as many of these industries conduct similar operations, a *Macro Industry* attribute was created by manually grouping industries from the Exiobase classification. Table 3 (Appendix A.3) displays the macro industry to which each specific industry is assigned.

For each entity, the attribute *Revenue* is calculated using the formula

$$Revenue = \frac{Total\ Environmental\ Cost}{Environmental\ Intensity\ (Sales)}$$

obtaining the revenues generated by each firm in a specific year, as Freiberg et al. (2021) notes that the *Sales* figures applied in the paper are actually firm revenue. Revenue is commonly used as a proxy for firm size, and can thus take the effects of firm size into account. As larger firms tend to have a bigger environmental footprint, dividing environmental impact by revenue allows for comparison between companies and therefore industries, regardless of individual size.

Thus, for each of the eight impact areas *(Abiotic Resources, Biodiversity, Working Capacity, etc.)*, a new attribute is calculated by evaluating the ratio with revenues. For example, the attribute *Rev_Abiotic* is calculated as

$$Rev\_Abiotic = \frac{Abiotic\ Resources}{Revenue}$$

These ratios represent the main dimensions that the dashboard is developed around.

### 4.1.3 ESG Data Extraction

The Python library yahooquery is utilized (Guthrie, n.d.) to extract the environmental part of ESG (Environmental, Social, and Governance) data from yahoo.finance.com. First, a query to the website is made to retrieve the ticker of each company in the data set. Then, ESG scores are scraped from the ticker's "sustainability page". The ESG data is provided by Sustainalytics, Inc, a specialist in evaluating the potential risks to a company's business value that may arise from ESG issues. These risks are assessed using a two-part framework that takes into account both the company's susceptibility to ESG issues specific to its industry and its ability to effectively address those issues. The resulting ESG Risk Ratings are rated on a scale from 0 to 100, with a lower score indicating lower unmanaged risk (Sustainalytics, n.d.).

### 4.1.4 Final Cleansing and Export

The final data set is created by combining the cleaned and filtered initial data set with the scraped ESG data. The data set obtained is then cleaned one last time by removing the columns that were only used for the merge. Finally, the data set is exported as an .xlsx file. In the following section, an overview of the final data set is provided.

## 4.2 Data Overview

The data set is composed of 4532 entries and 25 columns, described in Table 2 (A.2).

*Year* ranges from 2010 to 2019, and the occurrences are shown on the left side of Figure 1. 2018 has the most instances, with 544 entries, while 2010 has the least, with only 338. The data set contains 453 entries per year on average. The plot on the right-hand side of Figure 1 relates *Year* with *Environmental Intensity*; the figure shows how the environmental intensity is primarily concentrated between 0 and -0.5, with the vast majority of entries being negative, regardless of the year.

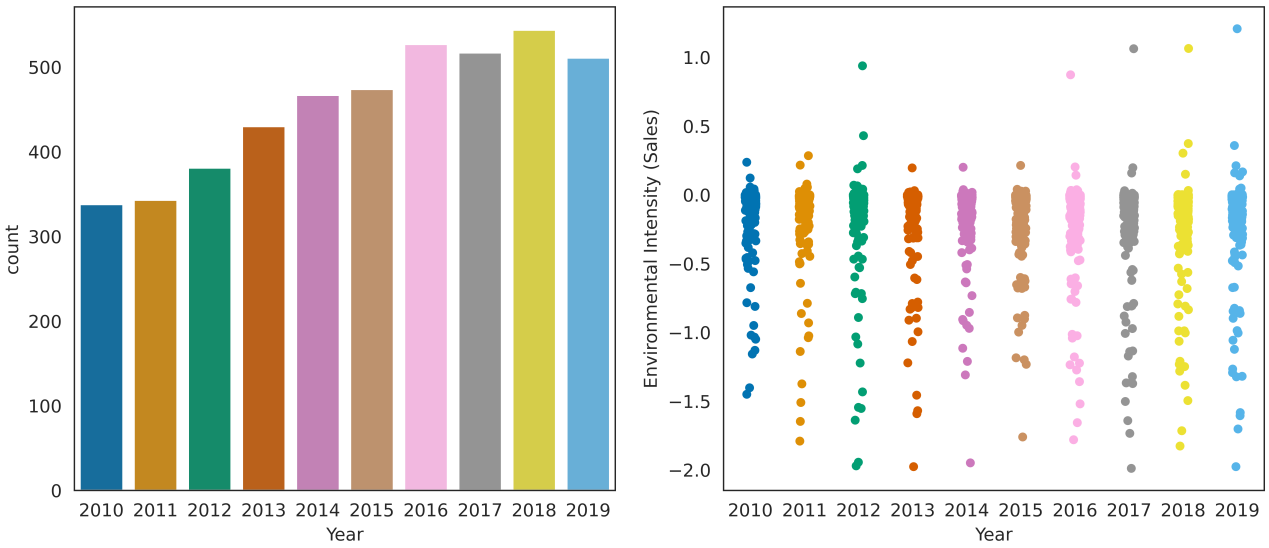


Figure 1: *Year* and *Environmental Intensity*

The *Macro Industry* classification described in section 4.1.2 yields a total of 15 industries, the occurrences of which in the data set are shown in Figure 2. *Finance & Insurance* has the most entities (781 entries) followed by *Real Estate & Construction* (593 entries) and *Tech & Communications* (555 entries). On average, each industry has around 300 entries.

The data set comprises 26 countries, with different frequencies. Figure 3 compares the frequencies of the 10 most common countries: at the top, the United Kingdom with 1608 entries, followed by France with 527 and Germany with 379.

Figure 4 shows the correlation between the numerical variables of the data set. As expected a positive correlation is present between some of the impact areas as the effects they capture

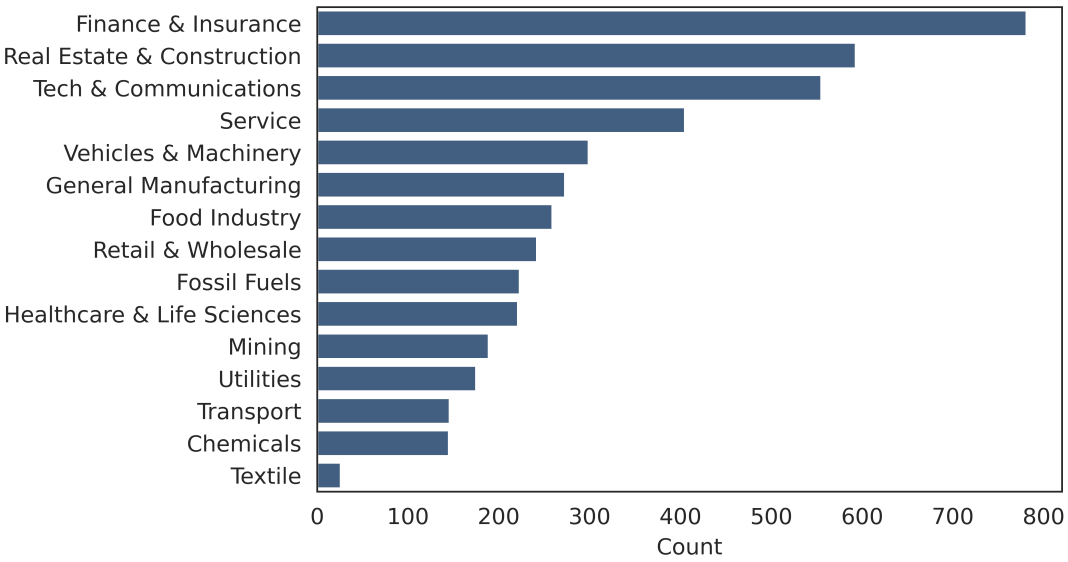


Figure 2: *Macro Industry*

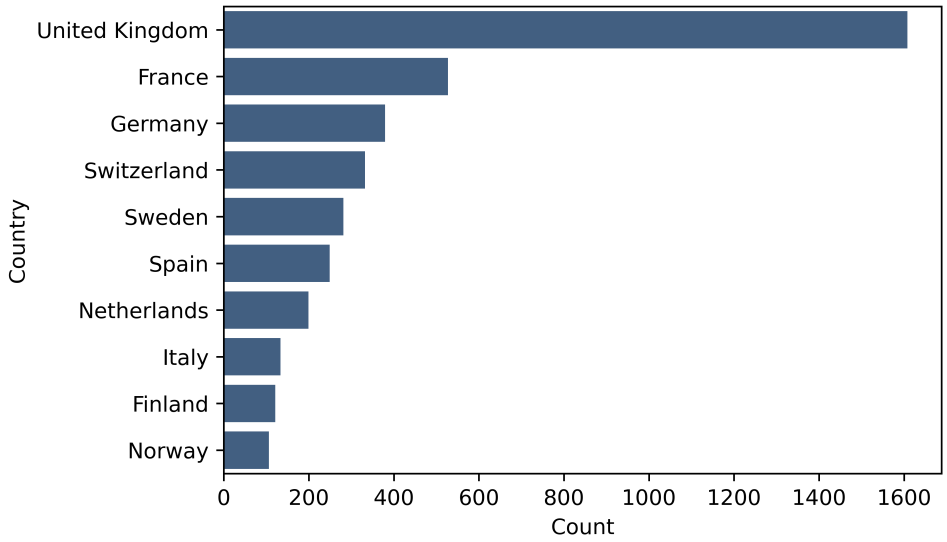


Figure 3: Frequencies of 10 most common countries

partially overlap. It is also interesting to notice that *Biodiversity* has close to none correlation with any other variable.

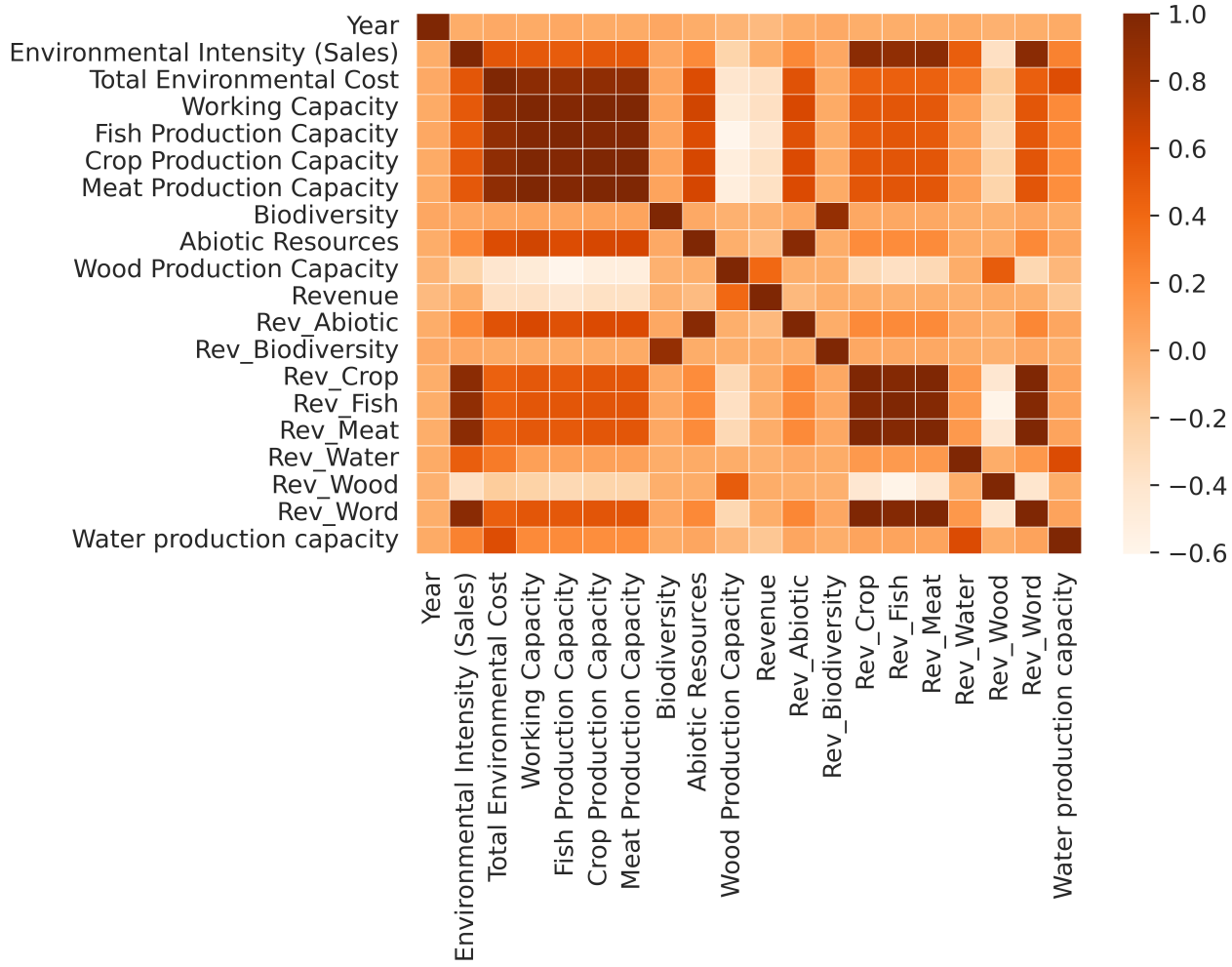


Figure 4: Correlation matrix

# 5 Dashboard

As the dashboard is meant to serve as a flexible tool for comparing and visualizing different types of environmental impact of European industries, its design must be intuitive and allow for in-depth exploration.

## 5.1 Design Choices

The dashboard incorporates descriptions, interactive visualizations to encourage user exploration, key performance indicators (KPIs), filters, and buttons. The dashboard objects are presented in a left-to-right sequence to guide the viewer through the narration, as the viewer's attention is initially drawn to the top left corner and center of the dashboard.

The selected impact area is highlighted in the upper left corner to ensure that the user is aware of their selection. Underneath, two KPIs showing the average Cost of Environmental Impact (CEI) and the Intensity of Environmental Impact (IEI) are displayed together with a trend line chart showing how the IEI changes over time. This provides the user with a high-level overview before going into more specific details. The trend line graph is chosen as it is effective at displaying long-term trends while also highlighting smaller changes. Comprehension is aided by colors that differentiate negative, detrimental values (displayed in a deeper hue) from positive, beneficial values (shown in a lighter color). Further, this graph allows the user to filter the entire dashboard by year.

The central, most prominent visualization is a map of the European continent, including only countries for which data is available. This map allows the user to instantly locate areas of interest, such as those with the greatest or least environmental impact as well as filter by clicking on a given country. When the user hovers over a country, that country's IEI over time is displayed with a line chart. This design allows the user to gain a more thorough understanding of the corporate impact on the chosen impact area of that country. As the hover instantly shows the national trend, it is simple to compare across countries.

At the bottom, the environmental intensity over time for each selected industry is shown. This graph helps to quickly identify which industries play a major role in the chosen impact area. Hovering over any point on this graph displays a tooltip with the year and average intensity and cost for that industry in that year. The title of each industry is displayed above the trend line graph, making it the most intuitive and efficient way of selecting a single industry.

Immediately beneath this, the average Environmental Risk Score from Sustainalytics' ESG Ratings is shown for each industry using a bar chart and an indicative color range with a deeper hue for worse values. As ESG data has become widely used in corporate reporting and sustainability analysis, it can be both used for further exploration of the issue at hand and as a sanity check for the main data, especially as the location allows for immediate comparison with IEI data.

### 5.1.1 Explanatory and navigational elements

The top of the dashboard contains descriptive information to give a quick overview to the user. This includes a descriptive title, a subtitle that demonstrates a generic question that the dashboard can answer, brief instructions and definitions, as well as a color legend. It also includes buttons for going back to the landing page, resetting filters, and opening the firm-level views of CEI and IEI. The firm-level view provides the individual scores for all entries included by the filters and is hidden by default to avoid information overload. These two buttons use the common *hamburger button* design, often used to indicate menus, to attract users to click on it, as well as to mimic the design of a table. The *Cost* button provides a visualization of the table containing the company's name, country, and CEI of the chosen impact area, while the *Intensity* button shows a similar table with IEI instead of CEI. The tables are displayed in ascending order, such that the companies with the greatest negative effects on each specific factor are shown first. When hovering over any company's score, the tooltip shows the industry and sub-industry above a table with year-specific data on IEI, CEI, and revenue. Revenue is included as a proxy for firm size, allowing for easy understanding of the firm's importance. Both CEI and revenue are denoted in millions of USD ($MM) to conserve space and for quicker comparison between the two.

The top of the dashboard is visually separated from the main part that includes all visualizations by a thin white line that emphasizes the split. On the right side, the user can filter by impact area and Industry. As impact areas are meant to be seen individually, these can be selected using a single-value list. On the other hand, it can be valuable to select multiple industries at the same for comparison, and as such, industries can be selected using a multiple-values list. As the contents of both the impact areas and industries can seem ambiguous, two info buttons placed near the filters provide definitions and explanations of the measurements used in IEI and CEI, as well as details on how the sub-categories are grouped.

### 5.1.2 Colors and Accessibility

As colors are highly effective instruments to raise the audience's attention (Knaflic, 2015), a meaningful choice of colors is necessary. To complement this selection, the light background results in a higher contrast with the elements of the dashboard and their colors, aiding those

with color-sight deficiencies. Three main aspects were considered for the choice of color palette:

- the readability and comprehension of the visualizations
- the semiological meaning of colors
- the accessibility issues that may arise

The theory of color-in-context suggests that the immediate perception of a color initiates evaluative processes that can impact psychological function. Therefore, it is important to consider the contextual setting when determining the meaning of a color. Maintaining a consistent palette ensures that the intended emotions are effectively conveyed to the user. (Schwarz, 2014)

As red is often associated with danger (Adobe, n.d.-c), threat, and caution (Schwarz, 2014), and as environmental pollution is a pressing concern, dark red (#943422) is used as an indicator for the lowest IEI values to evoke these emotions and effectively convey the gravity of the situation. On the other hand, yellow is often associated with caution, as is shown in the meaning of yellow road signs, and thus yellow (#EFCE5A) is appropriate to represent less severe negative values. IEI values that are very close to zero, both positive and negative along with zero itself, are represented in grey (#D7D4C7). Grey is often perceived as neutral as it is an achromatic color, meaning it is a color without hue. Finally, the values that are higher than zero, and represent a net positive impact, are represented by shades of light blue (#385E88) to indicate the positive nature. The automatic color palette by Tableau fits the requirements of the considerations well, as it automatically scales depending on the filtered in values, and aligns with these semiological meaning of colors once it has been reversed.

To ensure accessibility for all, the color palette is tested to ensure that the colors are easily distinguishable for those affected by any degree of color-blindness. The color palette's accessibility can be tested with respect to deuteranopia, protanopia and tritanopia through the Adobe Color Accessibility Instrument (Adobe, n.d.-b). No conflict is detected for the palette (Figure 7, A.4).

Furthermore, a high contrast ratio between the fonts and their background is used to make the text easy to distinguish. To ensure that the colors used for the text and background display a sufficient ratio of contrast, the "Adobe Color" contrast check (Adobe, n.d.-a) was employed.

The legibility of the text is examined by inputting each combination of font and background color. The check is passed by each combination of colors and font for standard text, larger text, and "icons and actionable graphics" for dark grey (#424141) (Figure 8, A.4).

Provided the dashboard's stated target audience and its use in the context of policy-making, along with its several functionalities and navigational elements that can be exploited only by using a computer, the dashboard is not optimized for smartphones and tablets, and is thus only configured for the *Generic Desktop* layout $1366 \times 768$.

# 6 Key Insights

As the main purpose of the dashboard is to be an exploratory tool that can provide a comprehensive overview of the environmental impact at the European level, the versatility of the tool makes it very well suited to function as a springboard into the most appropriate direction of further research. This versatility is shown by its ability to perform quick dives into any industry in any country, and decompose it into the specific effects of each company in each year. Due to the numerous combinations of characteristics that can be analyzed, a use case is provided to demonstrate the potential of the dashboard.

## 6.1 Use case: Water Production and Water Framework Directive

Water production is an important factor in the European Union (EU) in terms of both economic development and environmental sustainability. Water scarcity is a growing concern in many parts of the world, and the EU is no exception. As such, the EU has implemented a number of policies and initiatives to promote sustainable water management and production. One such initiative is the Water Framework Directive, which aims to protect and improve the quality of Europe's water resources. This directive establishes a framework for the protection of surface water and ground water, and sets out a number of objectives for the sustainable use of water resources. It also requires Member States to develop and implement measures to achieve these objectives, such as establishing water protection zones and improving the efficiency of water use. (European Commission, n.d.) Given that the Environmental Council could use the dashboard, they would have the possibility to make data-driven decisions on which industries and

countries are currently having the highest negative impact on water production and allocate funds accordingly to achieve the Water Framework goals.

By selecting the Water Production impact area, it is evident that the Average Environmental Intensity trend has significantly dropped after 2017. The countries with the worst scores are Portugal, Italy, and Greece, with Average Intensity scores of -0.06245, -0.05411, and -0.05342, respectively. The industry with the most detrimental impact on Water Production is the Utilities industry, which had an Average Intensity of -0.0784 and an Average Cost of -1,316,870,173 in 2019. The Fossil Fuels industry also had a high Environmental Risk score, followed by the Mining and Chemicals industries, indicating a potential for significant negative impact on the environment.

While the environmental intensity of the Water Production Area has improved since 2010, there is still a need for further efforts to address water scarcity and pollution. The dashboard can be used to conduct more detailed analysis of the reasons for the concerning scores of certain firms, countries, and industries, enabling the Environmental Council to develop targeted policies to achieve the goals of the Water Framework Directive.

It is important to note that the presented use case is a simplified example of a hypothetical real-world scenario and therefore has some limitations. It should not be viewed as a comprehensive representation of the work of a policymaker, but rather as an illustration of how the dashboard can be used to generate valuable insights when used as a starting point for further investigation. Despite these limitations, the use case demonstrates that the dashboard is a useful and user-friendly tool for addressing a variety of environmental policy-making issues.

# 7 Concluding Remarks

## 7.1 Limitations

One of the main limitations of our analysis is the uneven distribution of data from various countries, with some being significantly represented while others are either completely absent or have a limited number of data points. Another point of contention is the fact that the study is still a working paper, and has thus not been completed, nor peer-reviewed, further degrading

the current power of the data and dashboard. Further, the data set only contains information on a very small part of the total amount of firms in the European continent. Thus, it should be acknowledged that the insights generated by the dashboard may not provide a comprehensive view of the measures and industries.

If more data was available, the conclusions stemming from this dashboard would have more precise generalizability of the results, which may not accurately reflect the entire sample population. As of now, when filtering for specific countries, the size of the sample decreases such that detailed analyses on the industry level often cannot be performed, as there are a limited number of data points available for certain countries and certain industries. With a larger data set it would be possible to delve into the details of specific sub-industries and compare them in more depth (see Figure 6).

Additionally, the aggregation of industries in our analysis was based on our own understanding of which industries display similar traits, rather than using a standardized framework. Another limitation of our work is given by the fact that shaping environmental policies is a complex field that requires cross-sectoral know-how and analysis of all the possible externalities that may arise along with the interdependence of industries.

### 7.1.1 Weak use case - The example of Sweden and Biodiversity

As the data has several limitations, using it potentially leads to incorrect conclusions if not carefully examined. Therefore, the following presents a scenario in which the potential weaknesses in the data is more pronounced.

Consider the case of the Environmental Council tasked with allocating investments related to the European Union's Biodiversity Strategy for 2030 (Commission, n.d.). The council must determine which countries and industries have the most detrimental impact on biodiversity in order to effectively allocate funds. In order to accomplish this task, the following actions could be taken using our dashboard:

- Select the *Biodiversity* impact area filter
- Observe that Sweden is represented on the map visualization with a darker color, and upon hovering over it, observe that it has an Average Environmental Intensity score of

-0.0001800 and an Average Environmental Cost score of -1.963608

- Examine which industries have the most detrimental impact on biodiversity, specifically the General Manufacturing industry, which as of 2019 had an Average Environmental Intensity score of -0.0000054 and an Average Environmental Cost score of -93.231

While these results may be insightful at first, it is important to consider that the trends displayed in the data, specifically the large fluctuations, are influenced by the presence of one specific company in only certain years, and therefore may not be fully representative of the overall population. The disproportionate presence of entries from the Swedish company led to bias in the dataset, and therefore limited the reliability of the analysis.

## 7.2 Future Work

The dashboard can be improved in several ways. First, in the case of a significantly increased data set size, a closer analysis of sub-industries could prove insightful to see which sub-industries pollute the most. In the following tooltip, one can visualize just that. A policymaker in the Environmental Council, for example, may concentrate on the "Coal and Consumable Fuels" and "Copper" sub-industries to design policies aimed at minimizing their environmental footprint.

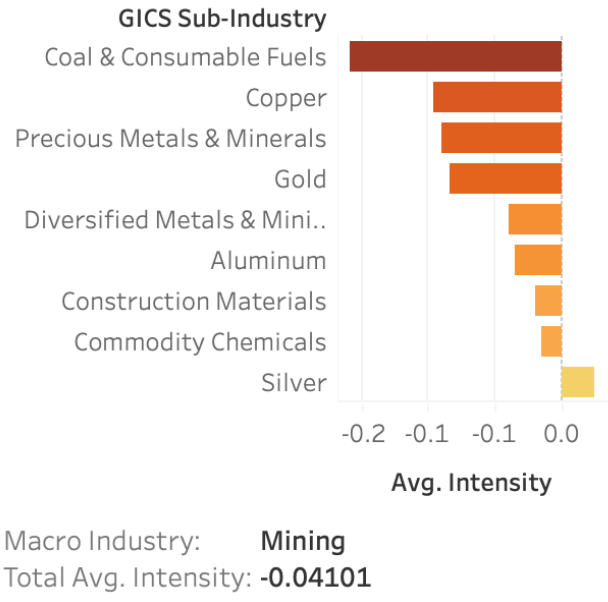


Figure 5: Subindustries with higher Environmental Impact

Several smaller features could improve the usability of the dashboard. First, the inclusion of a

filter button that allows the user to show only countries that are part of the EU on the dashboard would be advantageous for policy makers looking to address pollution and improve sustainability specifically within the EU, and without the influence of non-EU countries. Second, the Reset Filters button does not currently un-highlight the element that has been clicked by a user. This means that if a user has filtered by clicking inside a visualization, and then clicks the Reset Filters button, the filters will be reset, but the highlight stays. This is especially apparent when selecting a country. Lastly, as impact areas represent different areas of pollution, they are more powerfully seen either individually or as a whole. However, the areas are inherently related, and as such, the dashboard could offer a feature that allows for multiple areas to be selected at the same time.

# References


Adobe. (n.d.-a). Color Contrast Analyzer. https://color.adobe.com/create/color-contrast-analyzer

Adobe. (n.d.-b). Create color accessibility. https://color.adobe.com/create/color-accessibility

Adobe. (n.d.-c). What does red mean in design? https://www.adobe.com/creativecloud/design/hub/guides/meaning-of-red-in-design

Commission, E. (n.d.). Biodiversity Strategy Actions Tracker. https://dopa.jrc.ec.europa.eu/kcbd/actions-tracker/%5C#Financing%5C%20for%5C%20biodiversity

Consortium, E. (n.d.). Exiobase - About EXIOBASE. https://www.exiobase.eu/index.php/about-exiobase

et al., D. F. (2021). Corporate environmental impact: Measurement, data and information. *Harvard Business School*.

European Commission. (n.d.). Introduction to the EU Water Framework Directive. https://ec.europa.eu/environment/water/water-framework/info/intro_en.htm

European Council. (2022). Environment council configuration (env). https://www.consilium.europa.eu/en/council-eu/configurations/env

European Environment Agency. (n.d.). European Environment Agency. https://www.eea.europa.eu

European Parliament. (n.d.). Fact Sheets on the European Union - Environment policy: general principles and basic framework. https://www.europarl.europa.eu/factsheets/en/sheet/71/environment-policy-general-principles-and-basic-framework

Guthrie, D. (n.d.). yahooquery. https://yahooquery.dpguthrie.com

Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals. Wiley.

Schwarz, S. (2014). Red - take a closer look. *Plos One*. https://doi.org/https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0108111

Sustainalytics. (n.d.). ESG Risk Ratings. https://www.sustainalytics.com/esg-data

# A Appendices

## A.1 Tableau Dashboard Link

https://public.tableau.com/app/profile/freerik.g.tske/viz/PollutionDataonEuropeanIndustries_16729989409400/LandingPage

## A.2 List of data set attributes

| **Column Name** | **Explanation** |
|---|---|
| Year | year of the record |
| Company Name | name of the company |
| Country | fiscal residence of the company |
| Industry (Exiobase) | industry following Exiobase classification |
| Macro Industry | industry following macro classification |
| Environmental Intensity (Sales) | monetized environmental impact of the firm's operations during the specific year divided by revenue in that year |
| Total Environmental Cost | total monetized environmental impact of the firm's operations during the specific year |
| Revenue | revenue that the firm generated in the specific year |
| Environment Risk Score | environment risk score by Sustainalytics |
| Abiotic Resources | includes environmental impacts from mining |
| Biodiversity | captures land use and toxicity |
| Crop Production Capacity | captures soil degradation, air pollution, climate change, and land use |
| Fish Production Capacity | includes acidification, eutrophication, climate change, and toxicity |
| Meat Production Capacity | includes soil degradation, climate change, land use, and toxicity |
| Water Production Capacity | captures climate change, land use, and water pollution |
| Wood Production Capacity | includes climate change, air pollution, and land use |
| Working Capacity | captures human health effects from climate change, air pollution, and toxicity |
| Rev_Abiotic | abiotic resources on revenues ratio |
| Rev_Biodiversity | biodiversity on revenues ratio |
| Rev_Crop | crop production capacity on revenues ratio |
| Rev_Fish | fish production capacity on revenues ratio |
| Rev_Meat | meat production capacity on revenues ratio |
| Rev_Water | water production capacity on revenues ratio |
| Rev_Wood | wood production capacity on revenues ratio |
| Rev_Work | working capacity on revenues ratio |

Table 2: summary of data set attributes

## A.3 Exiobase Industries Grouping

Table 3 shows how 51 of the Exiobase industries have been grouped into 16 macro industries.

| |
|---|
| **Fossive Fuels** |
| Extraction of crude petroleum and services related to crude oil extraction, excluding surveying |
| Petroleum Refinery |
| Extraction of natural gas and services related to natural gas extraction, excluding surveying |
| Transport via pipelines |
| **Finance & Insurance** |
| Activities auxiliary to financial intermediation (67) |
| Financial intermediation, except insurance and pension funding (65) |
| Insurance and pension funding, except compulsory social security (66) |
| **Real Estate & ConstruEuropeanction** |
| Real estate activities (70) |
| Construction (45) |
| Transport |
| Air transport (62) |
| Other land transport |
| Sea and coastal water transport |
| Transport via railways |
| **Retail & Wholesale** |
| Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods (52) |
| Wholesale trade and commission trade, except of motor vehicles and motorcycles (51) |
| **Tech & Communications** |
| Computer and related activities (72) |
| Post and telecommunications (64) |
| Manufacture of office machinery and computers (30) |
| Manufacture of radio, television and communication equipment and apparatus (32) |
| Manufacture of electrical machinery and apparatus n.e.c. (31) |
| **Food** |
| Manufacture of beverages |
| N-fertiliser |
| Processing of Food products nec |
| Cultivation of cereal grains nec |
| **Mining** |
| Copper production |
| Mining of chemical and fertilizer minerals, production of salt, other mining and quarrying n.e.c. |
| Mining of other non-ferrous metal ores and concentrates |
| Mining of coal and lignite; extraction of peat (10) |
| Quarrying of sand and clay |
| **Healthcare & Life Sciences** |
| Health and social work (85) |
| Research and development (73) |
| Manufacture of medical, precision and optical instruments, watches and clocks (33) |
| **Service** |
| Other service activities (93) |

**continued from previous page**

| |
|---|
| Publishing, printing and reproduction of recorded media (22) |
| Education (80) |
| Recreational, cultural and sporting activities (92) |
| **General Manufacturing** |
| Manufacture of basic iron and steel and of ferro-alloys and first products thereof |
| Manufacture of fabricated metal products, except machinery and equipment (28) |
| Manufacture of rubber and plastic products (25) |
| Manufacture of tobacco products (16) |
| Paper, Forestry, logging and related service activities (02) |
| Manufacture of furniture; manufacturing n.e.c. (36) |
| **Textile** |
| Manufacture of textiles (17) |
| Manufacture of wearing apparel; dressing and dyeing of fur (18) |
| **Vehicles & machinery** |
| Manufacture of motor vehicles, trailers and semi-trailers (34) |
| Manufacture of machinery and equipment n.e.c. (29) |
| Sale, maintenance, repair of motor vehicles, motor vehicles parts, motorcycles, motor cycles parts and accessoiries |
| Renting of machinery and equipment without operator and of personal and household goods (71) |
| **Electricity & utilities** |
| Collection, purification and distribution of water (41) |
| Production of electricity by hydro |
| Production of electricity by petroleum and other oil derivatives |
| Production of electricity by solar photovoltaic |
| Production of electricity nec |
| **Chemicals** |
| Chemicals nec |

Table 3: Macro Industries Grouping

## A.4 Accessibility Figures

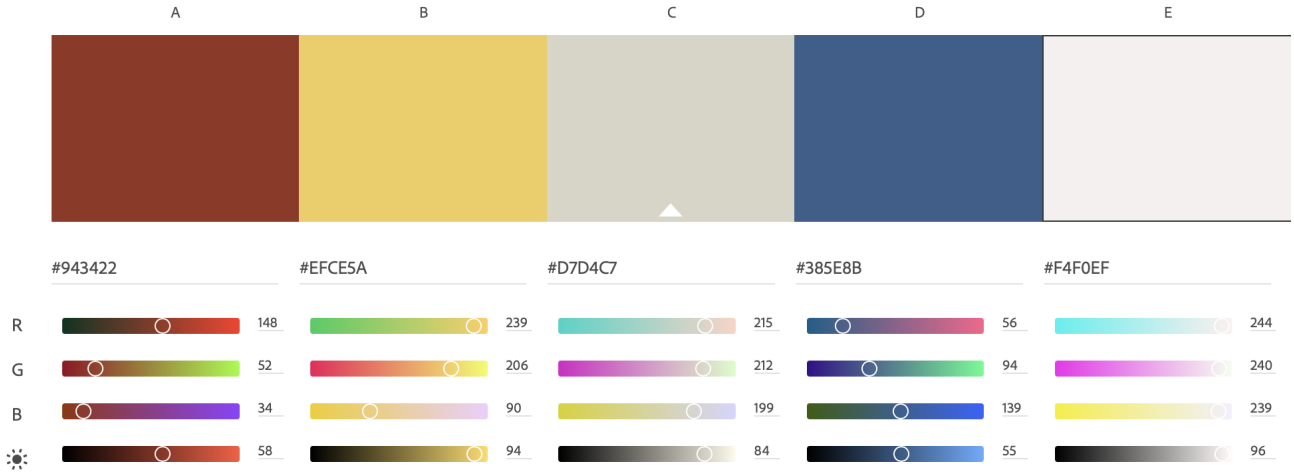


Figure 6: Color Palette (Adobe Color 2023)

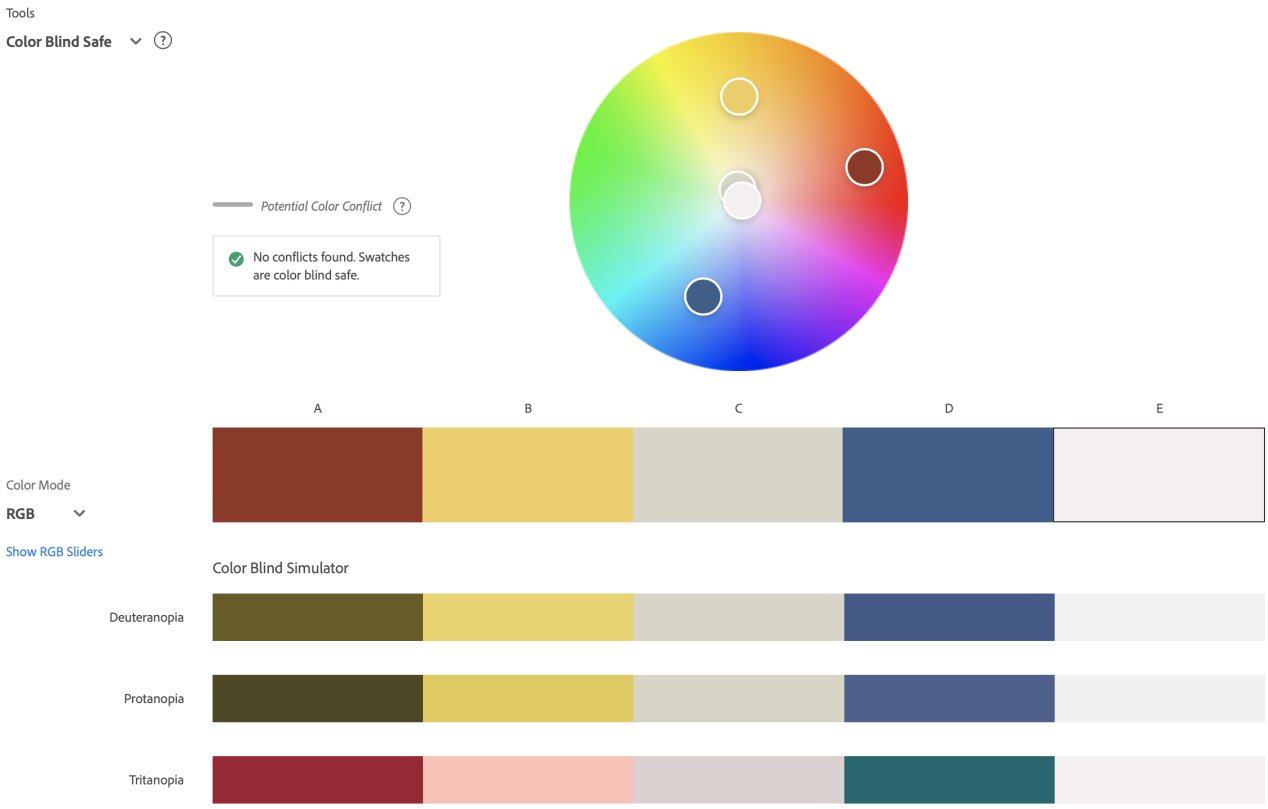


Figure 7: Accessibility Color Blindness Check (Adobe Color 2023)

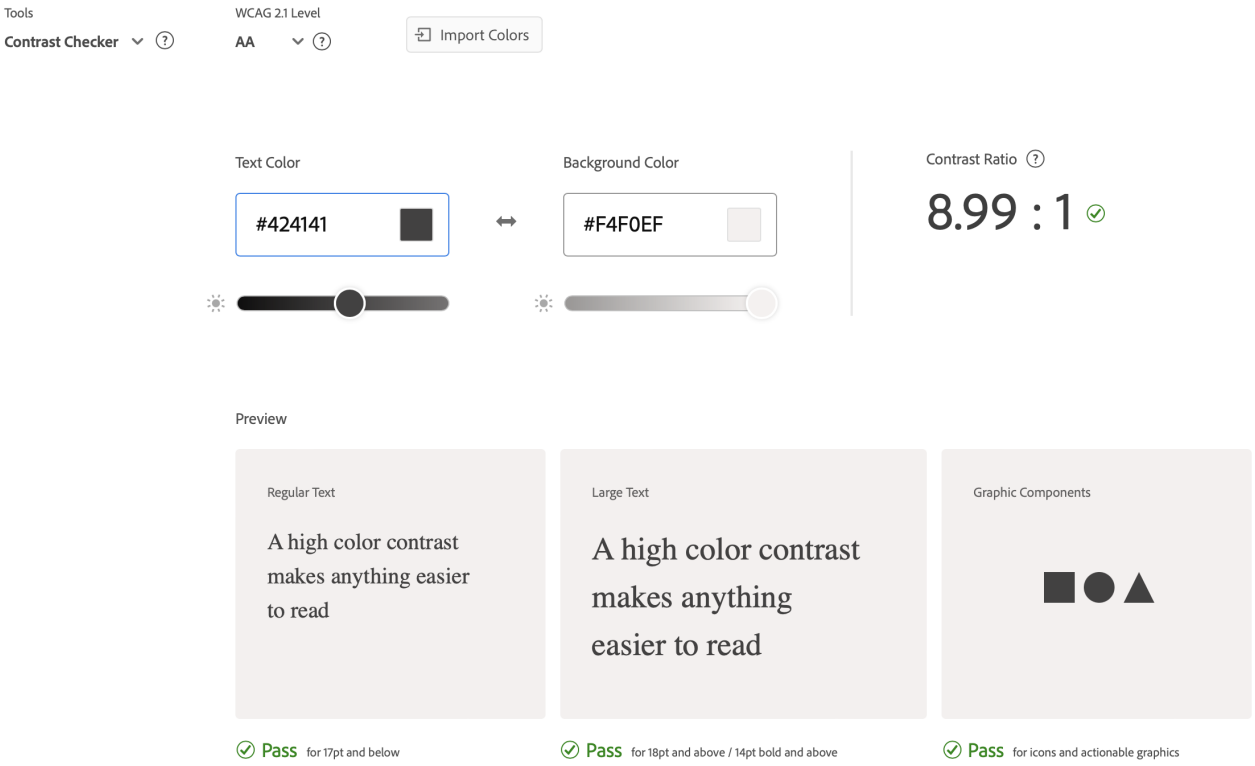


Figure 8: Accessibility Color Contrast Check (Adobe Color 2023)