**CCT College Dublin**

**Assessment Cover Page**

******

| **Module Title:** | **Strategic Thinking** |
| --- | --- |
| **Assessment Title:** | Ca2: Global plastic usage: future impact awareness |
| **Lecturer Name:** | **James Garza** |
| **Student Full Name:** | **Cristina Priolo** |
| **Student Number:** | Sba23037 |
|  |  |
|  |  |
|  |  |
|  |  |
| **Assessment Due Date:** | **17th December 2023** |
| **Date of Submission:** |  |

******

**Declaration**

| **By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.** |
| --- |

**Global plastic usage: future impact awareness**

**Table of content:**

[**Introduction: 4**](#_j3k0108k4zym)

[**Objectives: 4**](#_ijruzimrndqn)

[**Problem Definition: 4**](#_jccr8qhh7gnz)

[**Scope: 5**](#_tvmeyjlhs02j)

[*Table 1: Role and Responsibilities. 6*](#_w623ps2t0m83)

[*Table 2: Analysis of tasks 7*](#_f4153xmosvm2)

[*Table 3: Project planning 7*](#_7m80it2o5nz)

[**Potential data for the project: 8**](#_iiwvp5cl0c6u)

[*Table 4: Data sources 8*](#_uzkcpb7ek3eh)

[**‌Ethical considerations: 9**](#_ont3ewnbikqz)

[**Report Paper of Artefact: 10**](#_j0qphq5miiqz)

[Dataset analysis and exploration: 10](#_ko0o6ytzeakv)

[Table of missing values 12](#_uhedp1y9tbj)

[Graphic: Matplot of missing values 12](#_dtdfmn7fxhfc)

[Graphiac: HeatMap 13](#_3y98ej49z0c8)

[Graphic: Dendrogram 14](#_njnfhfn8jlem)

[Visualisation: 14](#_p6abw4vbmy4p)

[Pie chart of distribution of co 15](#_gbimc1w0009m)

[**References: 15**](#_3o0ajwaf15i)

# **Introduction:**

The main purpose of this report is to examine the usage of plastic and the global impact that it has to bring awareness to the readers.

Plastic pollution has emerged as one of the most pressing global environment challenges of our time. The report will delve into the complexities of the situation and its far-reaching significance, aiming to shed light on the myriad issues that underlie this crisis.

By examining the contributing factors of the increasing plastic usage, the economics of plastic production this report seeks to highlight a future forecast about the situation.

# 

# **Objectives:**

This assignment aims to explore the following objectives:

**1- Global usage of plastic:** Investigate and examine the current usage of plastic worldwide. Highlight main factors of consumption based on previous and actual data with mass production trends.

**2- Impact of plastic utilisation**: Forecast will indicate: What are the consequences? How will pollution and the ecosystem be?

**3- Waste management:** Exhibit the recycling data in order to respond to possible future risks.

By addressing these objectives, the goal is to raise awareness to have hypothetical sustainable solutions.

# **Problem Definition:**

This report will identify the key issues of global plastic usage. The global problem of plastic usage presents a multifaceted crisis that demands immediate attention. With each day passing, the excessive reliance on plastic deepend, leading to dire consequences for the environment and future generations.

Awareness is crucial for this report will delve into the gravity of the situation, emphasising the extreme need for immediate action. By addressing this crisis as early as possible this is essential to safe-guarding and preserving the planet.

Although, By addressing it holistically, this include:

1. Reducing single-use plastics
2. Recycling resources
3. Fostering global cooperation

Challenge of the project are the following:

* Appropriate data: finding dataset that have enough rows to support the objective of the analysis
* Avoid Bias in the analysis: staying neutral and focusing on the facts.
* Finding sustainable solutions for the problem supported with reliable dataset: the lack of insufficient data has led to continuous research. The project aims to raise awareness which gives the reader an opportunity to research sustainable solutions.

The context of the problem and the important of this to be addressed are:

*“Plastic pollution is a planetary threat, affecting nearly every marine and freshwater ecosystem globally. In response, multilevel mitigation strategies are being adopted but with a lack of quantitative assessment of how such strategies reduce plastic emissions The global threat from plastic pollution” (Borrelle et al. 2020, p.1).*

# **Scope:**

Over the two semester the scope of the project is to analyse the following topics and try to answer the following questions:

* Usage of plastic worldwide: which factors are impacting, mass production, forecast for future production of the plastic if nothing will change.
* Pollution and Ecosystem: How much the pollution increased and the ecosystem degraded? What are the causes?
* Analyse the recycling waste: what do we recycle? What is the capacity of the recycling facilities? Based on the analysis will we be able to plan and respond to the demand?

The project aims to bring at the end of the two semester attention on the topic to be more conscious about the long term effect on this concept.

Inclusions of the project:

* Definition of the problem and objective
* Analysis of the dataset of worldwide plastic usage
* Forecast on plastic usage
* Analysis of global pollution dataset
* Forecast about pollution
* Analysis of general waste
* Forecast about capacity of recycling facilities
* Conclusion of the analysis to bring awareness on the topic

Exclusions:

* Bias
* Avoid using personal data
* Not providing sustainable solution

Role and responsibilities:

| **Role** | **Who:** |
| --- | --- |
| Project Manager | Cristina/Hodan |
| Researcher | Cristina/Hodan |
| Data collection and cleaning | Cristina/Hodan |
| Data Analysis | Cristina/Hodan |
| Data visualisation | Cristina/Hodan |
| Writer | Cristina/Hodan |

## 

## *Table 1: Role and Responsibilities*.

Boundaries:

* Dataset limitation
* To avoid the limitation of the geographic area the project aims to analyse the situation on a global scale as to why the project title evolved during the time
* Research of solution to the argument
* Time frame of the project: Two semesters

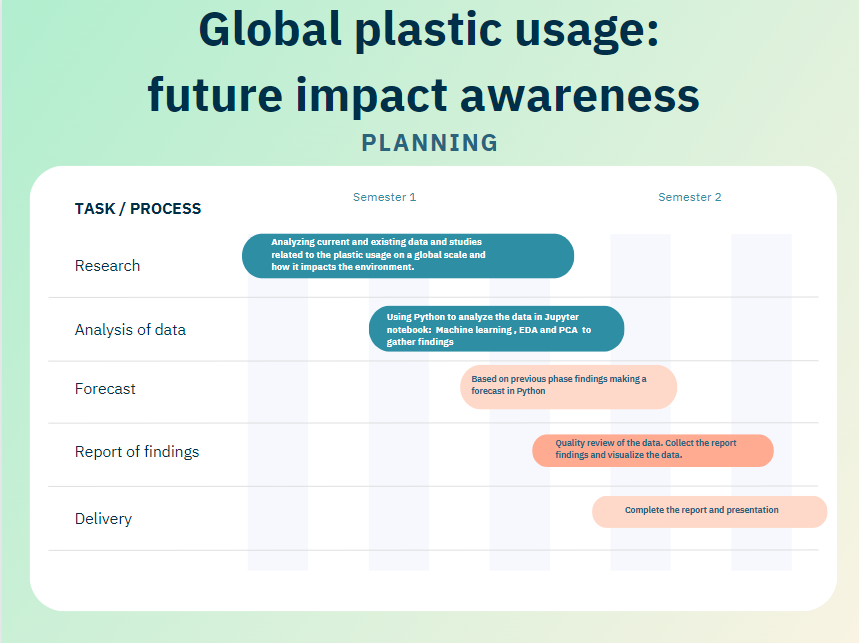
In-depth analysis:

|  | **First semester** | **Second semester** |
| --- | --- | --- |
| **Focus** | Foundational research: extensive literature review and data set collection | Analysis of garthing, the synthesis of findings and the formulation of rememendations |
| **Aims** | Prepare the material to move to the second semester phase. | Allow for a deeper exploration of global policies, case studies as well best practices and will also provide ample time for additional review if required which will ensure that the report quality. |

## 

## *Table 2: Analysis of tas*ks

Planning:



## 

## Table 3: Project planning

Project accomplishment:

The main recommendations and strategies for raising awareness, driving change as well as promoting recycling best practices. This approach will ensure a well rounded, evidence-based exploration of the global plastic usage issue and future awareness.

Final consideration on the project scope:

By the end of semester two we aim to deliver a comprehensive academic report that includes the following:

1. Extensive research: a well-researched report featuring a thorough literature review, primary and secondary research, data analysis and an in depth exploration of global plastic usage. This will include:

* Daily plastic consumption
* Global pollution
* Recycling waste facilities
* Forecast based on the data

1. Data-driven recommendations: This will include evidence based recommendations for raising awareness to mitigate the future impact of plastic usage

# **Potential data for the project:**

| **Data Source** | **Data amount** | **Permission** |
| --- | --- | --- |
| www.kaggle.com. (n.d.). Global Plastic Pollution. [online] Available at: <https://www.kaggle.com/datasets/sohamgade/plastic-datasets> | Full | open resources allowed by their terms and conditions |
| Datopian (n.d.). Daily\_csv plastic monkey 78. [online] DataHub. Available at: https://www.datahub.io/gitchenze/daily\_csv-plastic-monkey-78 [Accessed 15 Oct. 2023] | Full | open resources allowed by their terms and conditions |
| Our World in Data. (n.d.). Extrapolated change in plastic fate. [online] Available at: <https://ourworldindata.org/grapher/plastic-fate-to-2050>. | Full | open resources allowed by their terms and conditions |

## 

## Table 4: Data sources

The data that has been found as potential for the project are from open resources so they are allowed to be used by their terms and conditions.

# ‌**Ethical considerations:**

While the report global plastic usage: future impact awareness does not directly involve sensitive data, user privacy or potential societal impact, however we believe ethical considerations remain essential to the report.

The report will prioritise transparency and accuracy in the use of data, adhering to proper dataset usage.

All the sources referenced will be appropriately cited under the guidelines of Harvard Reference to acknowledge the contribution of others to avoid plagiarism.

Additionally the report will emphasise the importance of responsible data handling and the ethical use of information for academic and educational purposes.

# **Report Paper of Artefact:**

The artefact wants to research and analyse the dataset “per-capita-plastic-waste-vs-gdp-per-capita” to address the problem of plastic consumption on a global scale.

The report examines the variables and the data for researching implications and forecasting future possible scenarios.

By exploring historical data, investigating the geographic and temporal aspect of the chosen dataset, and employing data analysis techniques, this report aims to bring awareness to the reader to start considering sustainable solutions to mitigate the impending crisis. In the optic to data analysis techniques.

This part will cover:

* **Dataset analysis and exploration**: discussion of the findings and insights to drawn a preliminary preliminary analysis of the dataset to see the next steps
* **Visualisation techniques:** application of visual representations to illustrate trends, patterns, and relationships within the dataset.
* **Data Processing and Visualization**: explanation of how the data was collected and processed:
  + Source of data, data cleaning methods used, handling missing values, and encoding categorical variables.
  + Methodology explanation behind the choices
* **Model discussion**: discussion of the models implemented:
  + Explain the machine learning models used and reason behind the selection.
  + Describe the model training process, parameter tuning, and validation techniques used.
* **Conclusion**: Analyse the findings obtained from the models and data analysis:
  + Interpretation of the results obtained from the models in the context of the domain of plastic usage.
  + Conclusions from the analysis and discussion of the implications.
  + Insights into the significance of the findings and their relevance to the project's objectives.

# 

## **Dataset analysis and exploration**:

It has been chosen to use the public dataset “Global Plastic Pollution” from Keggle. The dataset has multiple csv inside so previous it was a preliminary analysis to use the one that was most useful for the current analysis. Between all it has been chosen the one per-capita-plastic-waste-vs-gdp-per-capita.csv. The reason behind was that it was most suitable for the purpose and had enough features and observations.

To do so the process of the dataset analysis and exploration is important to have a view over the raw data and the pattern in it. This is a vital step to extract important information and/or details from it.

With the help of the libraries the dataset. To start analysing the dataset the import of the libraries comes to help:

* Pandas: This library is used for data manipulation and analysis that help speed the analysis
* Matplotlib: a plotting library to create visualisation in Python
* Numpy: Helps for numerical computation and working with arrays
* Seaborn: statistical library for statistical visualisation

Once the dataset is imported with the proper function = pd.read\_csv the dataset can be analysed in its raw data. The chosen dataset has 48168 rows and 7 columns, so this means that the dataset has: 48168 observations and 7 features.

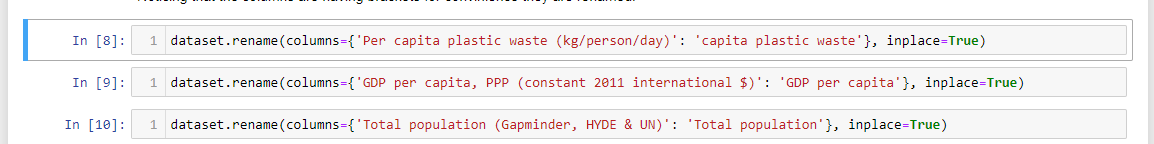
It is important to see what type of variables are in the dataset. This step is necessary to see how to treat the variable of the dataset for the next coming steps. The function dataset.dtypes shows the following:

| **Feature** | **Type** |
| --- | --- |
| Entity | object |
| Code | object |
| Year | int64 |
| Per capita plastic waste (kg/person/day) | float64 |
| GDP per capita, PPP (constant 2011 international $) | float64 |
| Total population (Gapminder, HYDE & UN) | float64 |
| Continent | object |

From this it is clear that what is an integer is a numeric data Numeric data type for numbers without fractions. A float variable is a numeric data type for numbers with fractions; an object is a mix of variables like strings and numbers.

This is vital to understand how to proceed next when the missing values will be handled.

The columns name are having brackets so for convenience they have been renamed with the rename function to avoid future problem in calling this columns:



It is crucial to start checking the dataset for duplicates and missing values in order to proceed with the next steps: visualisation and data processing.

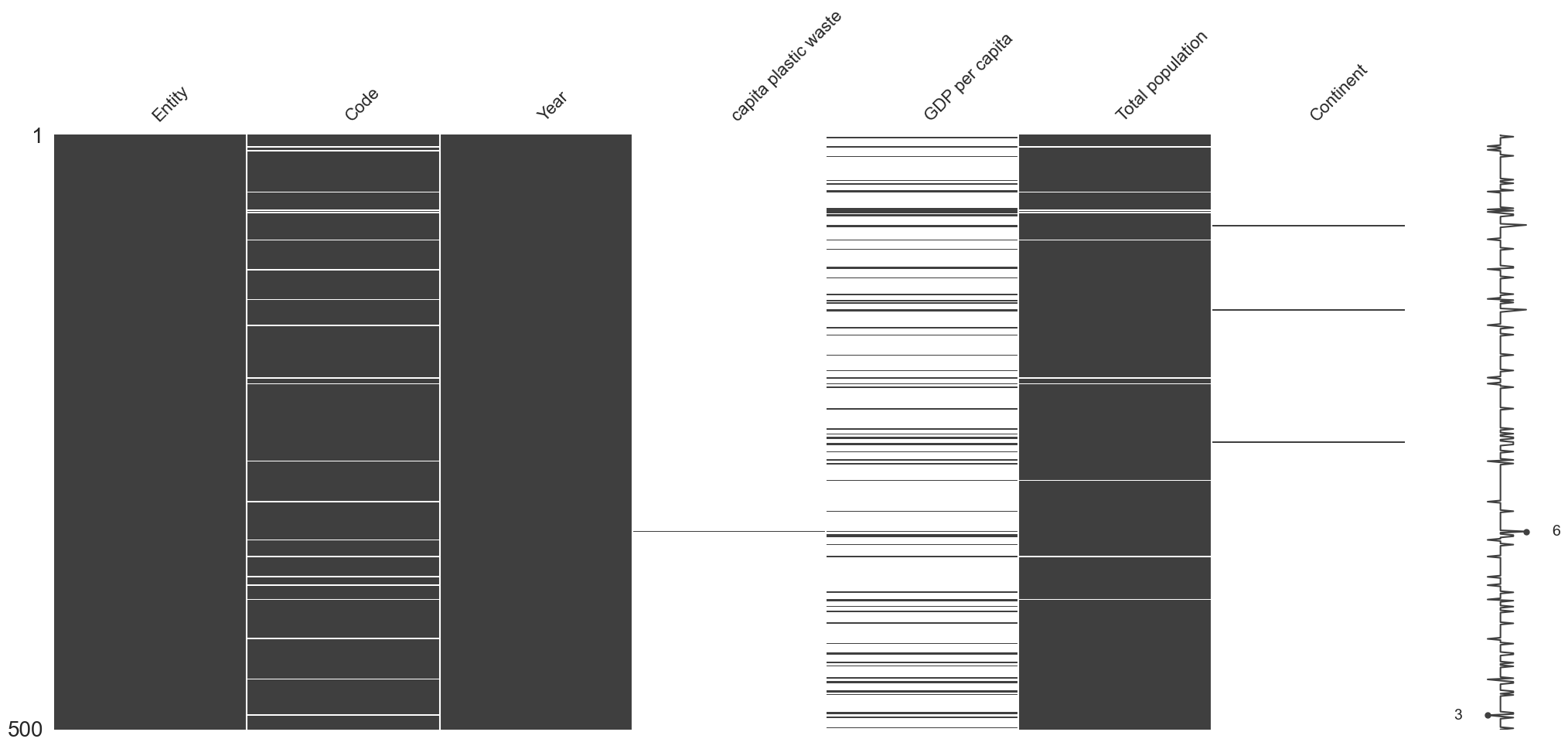
The dataset has 0 duplicated values but it has the following missing values for a total of 140925:

| **Feature** | **Amount of missing Values** |
| --- | --- |
| Entity | 0 |
| Code | 2014 |
| Year | 0 |
| capita plastic waste | 47982 |
| GDP per capita | 41761 |
| Total population | 1285 |
| Continent | 47883 |

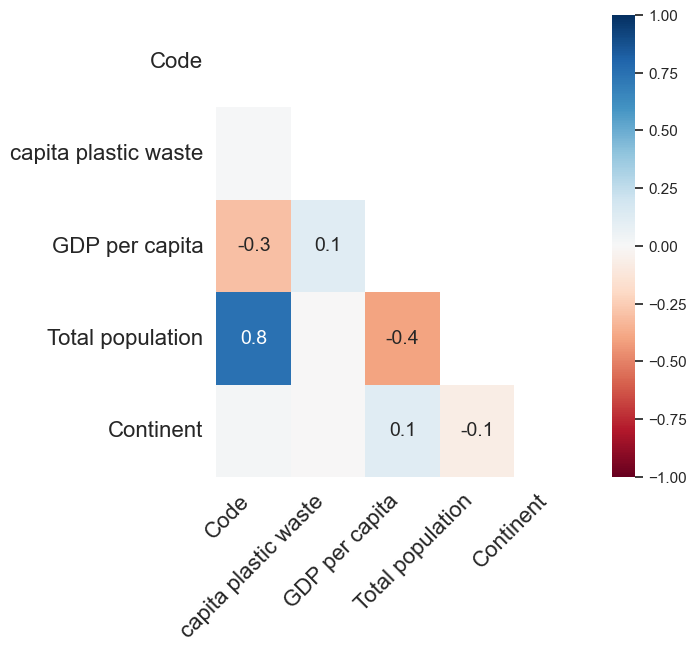
### Table of missing values

The library Missingno helps in the visualisation of the missing values.

The matrix identify the location of the missing data within the dataset:

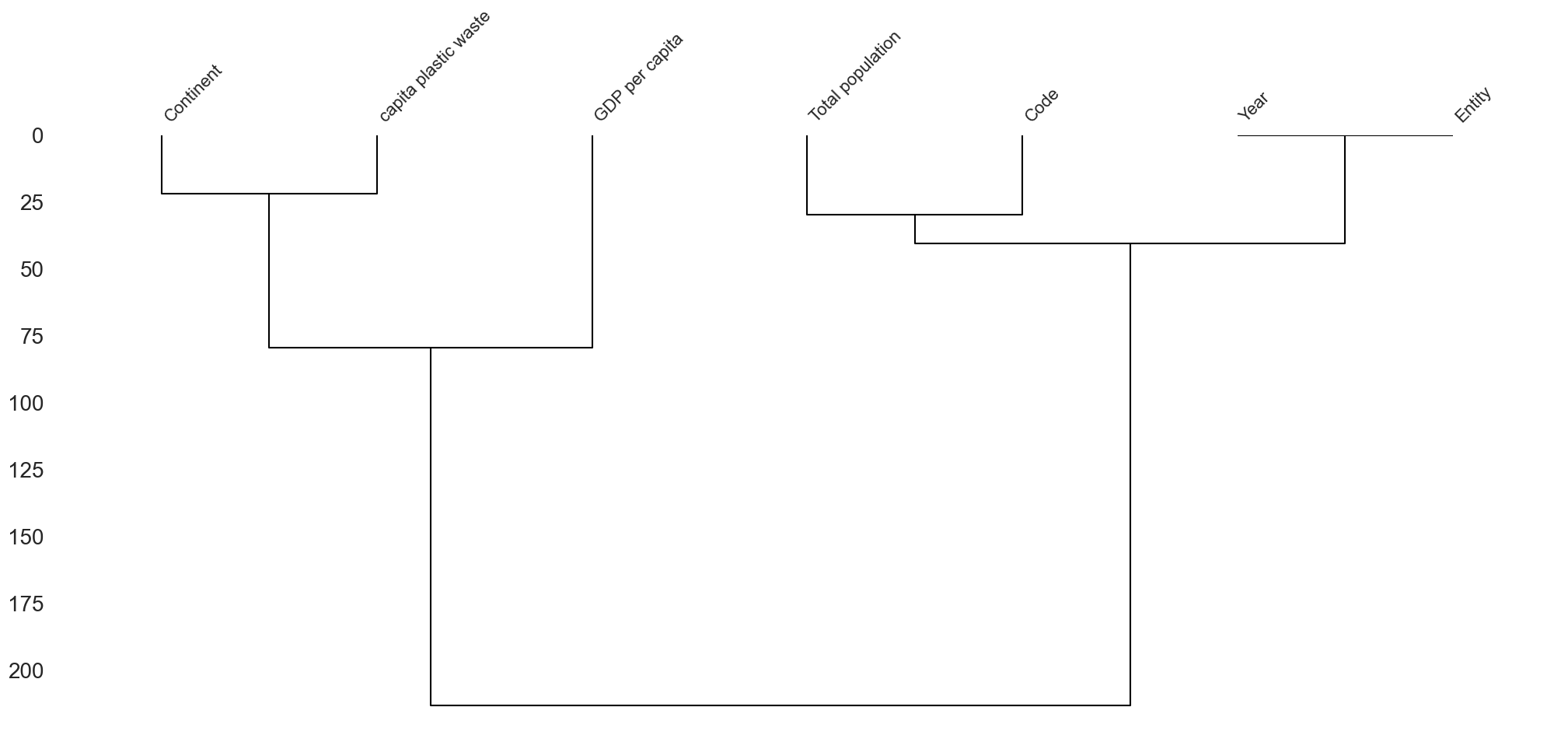


### *Graphic: Matplot of missing values*

The heatmap shows the correlation between missing values across different columns. This help understand the relationship inside the missing values inside different columns, if there is any: 

### *Graphiac: HeatMap*

The dendrogram identify the group of columns that have similar patterns of the missing values:



### *Graphic: Dendrogram*

## **Visualisation**:

Before preceding in handling the missing values to see the pattern and understand the dataset it is performed some visualisation:

### *Pie chart of distribution of co*

# 

# **References**:

1. www.kaggle.com. (n.d.). Global Plastic Pollution. [online] Available at: <https://www.kaggle.com/datasets/sohamgade/plastic-datasets>.
2. Datopian (n.d.). Daily\_csv plastic monkey 78. [online] DataHub. Available at: https://www.datahub.io/gitchenze/daily\_csv-plastic-monkey-78 [Accessed 15 Oct. 2023].
3. Our World in Data. (n.d.). Extrapolated change in plastic fate. [online] Available at: <https://ourworldindata.org/grapher/plastic-fate-to-2050>.
4. Oa, A. (2019). Public and Environmental Health Effects of Plastic Wastes Disposal: A Review. clinmedjournals.org, [online] 5(1). doi:<https://doi.org/10.23937/2572-4061.1510021>.
5. Borrelle, Stephanie B., et al. “Predicted Growth in Plastic Waste Exceeds Efforts to Mitigate Plastic Pollution.” Science, vol. 369, no. 6510, 18 Sept. 2020, pp. 1515–1518, <https://doi.org/10.1126/science.aba3656>
6. [Table 1: Role and Responsibilities](#_w623ps2t0m83)
7. [Table 2: Analysis of tasks](#_f4153xmosvm2)
8. [Table 3: Project planning](#_7m80it2o5nz)
9. [Table 4: Data sources](#_uzkcpb7ek3eh)
10. Avinash Navlani. “Random Forests Classifiers in Python.” DataCamp Community, 2018, www.datacamp.com/community/tutorials/random-forests-classifier-python.
11. Avinash Navlani. “Decision Tree Classification in Python.” DataCamp Community, 2018, www.datacamp.com/community/tutorials/decision-tree-classification-python.
12. javaTpoint. “Machine Learning Decision Tree Classification Algorithm - Javatpoint.” Www.javatpoint.com, 2021, www.javatpoint.com/machine-learning-decision-tree-classification-algorithm.
13. <https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm>
14. Jaiswal, Abhishek Sheshnath . Importance of Data Exploration in Data Analysis a Review Paper. 2 Apr. 2022, p. 1.
15. “Pyplot Tutorial — Matplotlib 2.0.2 Documentation.” Matplotlib.org, matplotlib.org/2.0.2/users/pyplot\_tutorial.html.
16. “What Are Data Types and Why Are They Important?” Amplitude, amplitude.com/blog/data-types#integer-int. Accessed 11 Dec. 2023.