FIN41910 Green Data Science

Group Assignment Title:
Analysis of corporate emissions drivers

Presented by

- Parary, Aman - 18417714
- Nagar, Shraddha - 22200477
- Kuntoji, Rohan -22202093
- Fernandes, William - 22201553

Executive summary

This report presents an analysis of the relationship between corporate financial performance and carbon emissions. The objective of the study is to explore the relationship between carbon emission and financial factors, and uncover opportunities for sustainable practices. The analysis employs regression models to test the relationships between independent financial variables and carbon emissions.

On the basis of literature review, the study formulates several hypotheses and develops regression models to test these hypotheses. The independent variables considered include efficiency ratio, liquidity ratio, valuation ratio, coverage ratio, profitability ratio, leverage ratio, and cost of capital. These variables are used to explain the dependent variable, carbon emissions. The data used in the analysis comprises carbon emissions data obtained from the Race to Zero data explorer and financial performance indicators fetched from Refinitive.

Exploratory data analysis (EDA) is conducted to uncover relationships and patterns between variables. The distribution of carbon emissions across sectors is examined, correlations across sectors are visualized, and trends in Scope 1 and Scope 2 emissions over time and across sectors are analysed. The panel OLS regression technique is then employed to estimate five different regression models using the dataset.

The empirical results section presents the findings of the exploratory data analysis, including sector-wise distribution of companies, summary statistics of emission-related variables, identification of outliers and skewness through box plots, and correlations between carbon emissions and financial variables visualized through heat maps. Furthermore, it discusses the mean sector performance in carbon emissions and the observed trends over the years. The panel OLS regression results provide an overview of the regression models and their statistical significance. The findings suggest that specific financial performance indicators have a significant influence on carbon emissions.

Introduction

Overly high carbon emissions are becoming a major cause of environmental deterioration and a barrier to sustainable growth. The achievement of business sustainability is largely dependent on corporate carbon performance. Promoting carbon performance requires a clear understanding of the variables that affect business carbon emissions. The primary source of climate change is still carbon emissions. Given the catastrophic effects of climate change and the enormous pace of greenhouse gas emissions, regulating carbon emissions reactively is no longer viable. (Oyewo, 2023). Carbon emissions that come from sources within the control or ownership of an organization are referred to as scope 1 emissions. Examples of these emissions are those brought on by the burning of fuel in boilers, furnaces, and cars. Indirect emissions under Scope 2 are those brought on by the purchase of energy, steam, heat, or cooling. Scope 2 emissions are accounted for in a company's inventory even though they are really produced at the plant where they are produced due to the organization's energy consumption. The aim of the analysis is to analyze financial data and carbon emissions to understand the relationship between them, identify financial factors influencing carbon emissions, and explore opportunities for sustainable practices.

Literature review and hypothesis development:

The relationship between an entity's financial performance and emissions performance has generally been explained by a number of hypotheses. The amount of time and money spent on emissions reduction solutions is small for low-carbon intensive businesses. For these companies, switching to cleaner manufacturing processes from end-of-pipe treatments results in increased Corporate Environmental Performance (CEP) and greater financial advantages. Due to increased efficiency and less waste, these businesses also become more environmentally sensitive, which lowers transaction costs connected with pollution abatement. This results in enhanced economic advantages (Ogunrinde, Shittu and Dhanda, 2022). Based on a sample of 652 American industrial companies, (King and Lenox, 2001) discovered evidence of a positive association between reduced pollution levels and better financial performance.

Based on the review of existing literature the following variables were chosen as independent variables to explain the dependent variable – Carbon Emissions (Scope 1, Scope 2, and their sum).

Efficiency Ratio - Asset Turnover:

H₁ : Asset turnover may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Asset\ Turnover_{i,t-1} + \beta_j * \begin{pmatrix} Current\ Ratio_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Liquidity Ratio - Current Ratio:

H₂ : Liquidity may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Current\ Ratio_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Valuation Ratio – EV to EBIT:

H₃ : EV to EBIT may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * EV\ to\ EBIT_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Coverage Ratio - Interest Coverage Ratio:

H$_4$: Interest coverage ratio may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Interest\ Coverage_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Valuation Ratio - Price to Cash Flow per Share:

H$_5$: Higher price to cash flow ratio may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Profitability Ratio - Return on Assets:

H$_6$: Profitability may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Return\ on\ Assets_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Leverage Ratio - Total Debt to Total Equity:

H$_7$: Leverage may be negatively associated with carbon emissions

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * Total\ Debt\ to\ Total\ Equity_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ WACC_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Cost of Capital - Weighted Average Cost of Capital (WACC):

H8 : WACC may be negatively associated with carbon emissions.

$$\log(Carbon\ Emissions_{i,t}) = \beta_0 + \beta_1 * WACC_{i,t-1} + \beta_j * \begin{pmatrix} Asset\ Turnover_{i,t-1} \\ Current\ Ratio_{i,t-1} \\ Interest\ Coverage_{i,t-1} \\ EV\ to\ EBIT_{i,t-1} \\ Price\ to\ Cash\ Flow\ per\ Share_{i,t-1} \\ Return\ on\ Assets_{i,t-1} \\ Total\ Debt\ to\ Total\ Equity_{i,t-1} \end{pmatrix} + \varepsilon_{i,t-1}$$

Sample and Data

The data related to carbon emissions was web-scrapped from the Race to Zero data explorer which is created by Climate Arc with data provided by CDP and SDG Labs. The financial performance indicators were fetched from Refinitive. The data from both the sources was cleaned and organised. After checking the data from the defined universe of 500 companies for any missing values, the final data comprised of 128 companies for a period of four years. To reduce the impact of extreme values and address non-normality and heteroscedasticity, natural logarithm transformation was applied to the carbon emissions data (Konadu et al., 2022).

Methodology

The 128 companies were analysed to check their distribution as per the Global Industry Classification Standard (GICS) sector classification. Exploratory Data Analysis was conducted to uncover relationships, patterns, and similarities between variables. Analysis included checking the distribution of carbon emission across sectors for the 128 companies, visualising the correlations across sectors and the sector performance on the carbon emission metric overtime. Further the carbon emission data was analysed to establish the Scope 1 and Scop 2 emission trends over time and across sectors.

The panel OLS regression technique was employed to estimate five different models (Model 1, Model 2, Model 3, Model 4, and Model 5) using a dataset consisting of 512 observations, to analyse the relationship between the independent and dependent variable.

Empirical Results

Exploratory data analysis (EDA) was conducted to visualize the number of companies per sector in the dataset. The results showed highest number of companies in the industrial sector, followed by consumer staples.

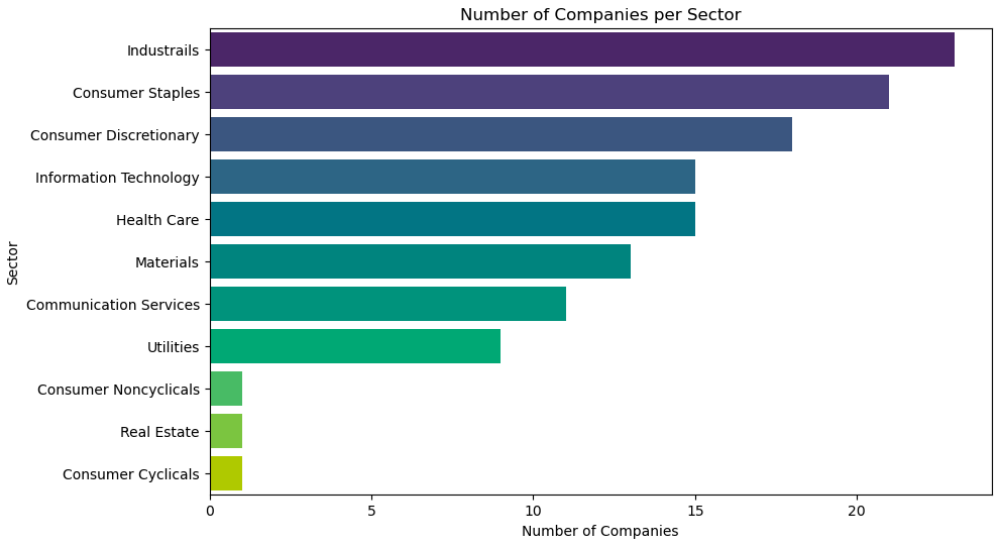


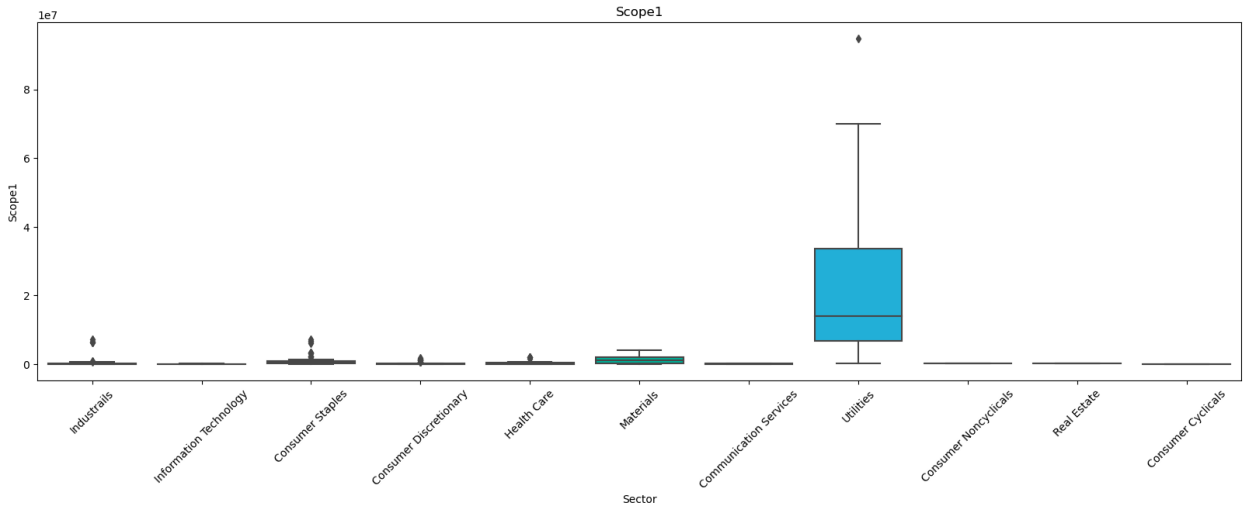
Table1 provides summary statistics for emission-related variables, including Scope1 and Scope2 emissions. The mean Scope1 emission value is approximately 2,069,802 tonnes of carbon dioxide equivalent, indicating the average level of direct emissions. The standard deviation of 8,137,275 tonnes of carbon dioxide equivalent suggests a significant variation in Scope1 emissions across the dataset. The minimum value of 28 tonnes of carbon dioxide equivalent and the maximum value of 94,794,693 tonnes of carbon dioxide equivalent demonstrate the range within which Scope1 emissions are distributed.

The summary statistics of financial variables such the Asset Turnover variable has a mean of 0.8, indicating the average efficiency of using assets to generate revenue. These summary statistics provide insights into various financial and performance metrics, helping us understand the average levels and characteristics of these variables in the dataset.

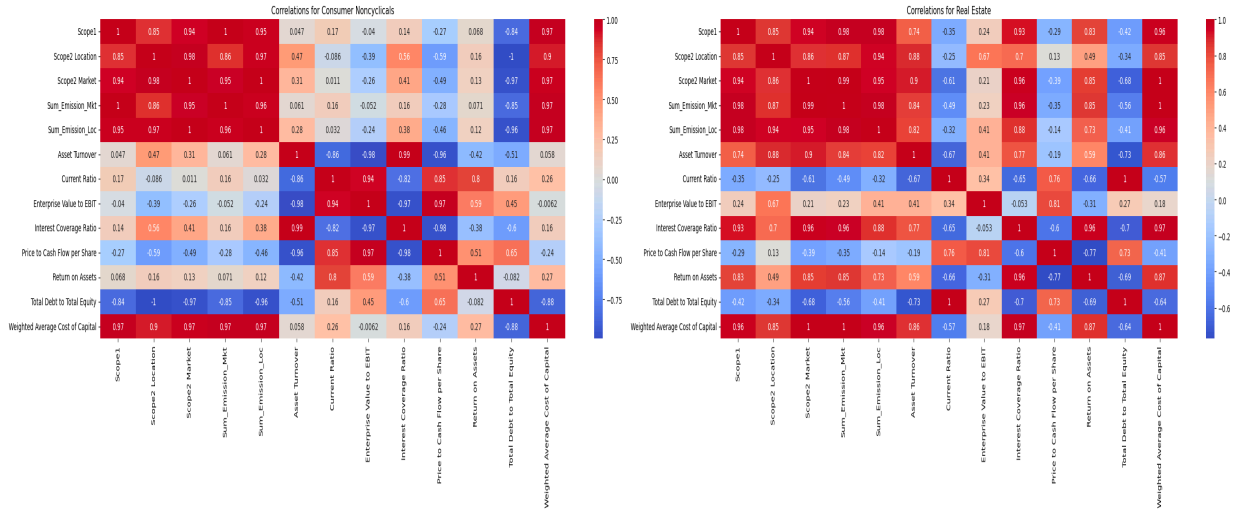
Table1: Summary Statistics

	Scope1	Scope2 Location	Scope2 Market	Sum Emission Mkt	Sum Emission Loc	Asset Turnover	Current Ratio	Enterprise Value to EBIT	Interest Coverage Ratio	Price to Cash Flow per Share	Return on Assets	Total Debt to Total Equity	Weighted Average Cost of Capital
count	512	512	512	512	512	512	512	512	512	512	512	512	512
mean	20,69,802	7,47,378	5,34,023	26,03,825	28,17,181	0.80	1.42	28.49	51.51	14.34	7.36	100.87	5.87
std	81,37,275	14,67,282	11,90,225	87,05,553	87,01,119	0.51	0.71	72.50	313.06	9.68	5.20	123.27	2.05
min	28	2,347	61	2,211	4,055	0.08	0.36	4.48	-5.29	1.36	-9.44	0.14	1.05
25%	25,536	62,835	28,175	91,849	1,28,669	0.46	0.94	12.78	5.62	7.33	4.27	44.32	4.42
50%	1,55,940	2,67,404	1,55,883	3,37,117	4,47,344	0.67	1.26	16.97	9.87	12.45	6.50	71.05	5.76
75%	6,08,561	7,30,836	5,02,587	12,10,011	15,92,783	0.98	1.70	22.64	24.07	19.07	9.02	126.85	7.38
max	9,47,94,693	1,30,57,352	1,20,22,083	9,64,15,272	9,58,84,305	3.30	5.58	937.94	5,073.53	75.02	40.06	2,032.34	13.28

EDA also included box plots to determine the outliers and resulting skewness in the data. Below plot shows positive skewness in the Scope 1 emission data for the utilities sector.

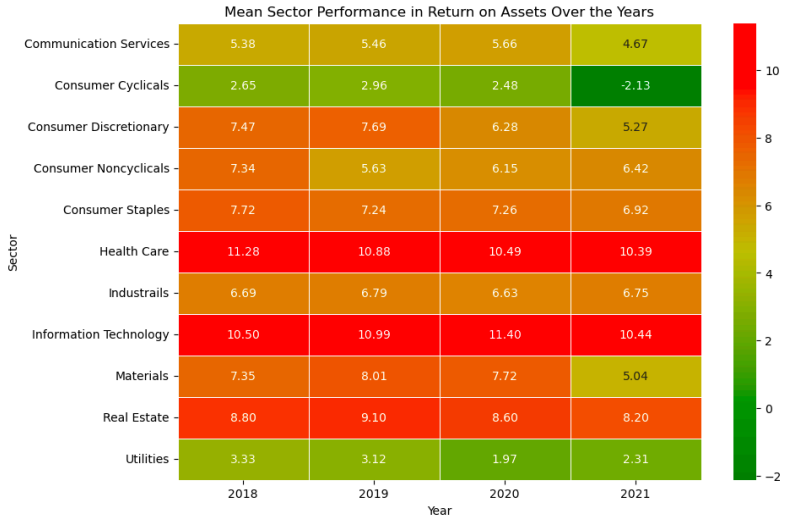


Heat map of correlations for each sector were plotted and anlysed. Such heat maps were created for all sectors using all emission and financial variables.

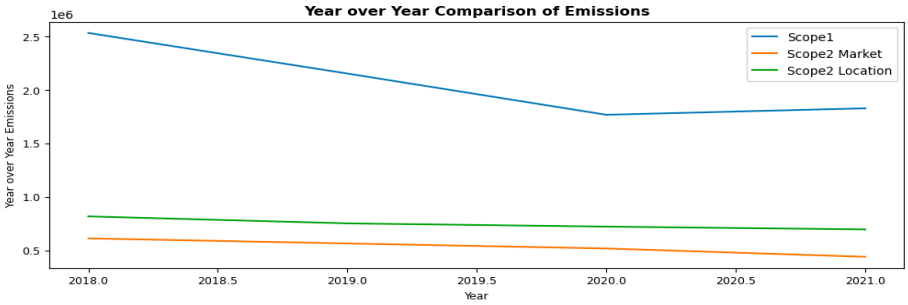


The analysis of mean sector performance in Scope 1 emission over the years showed highest emissions in absolute terms for the utilities sector with a decreasing trend in the following years.

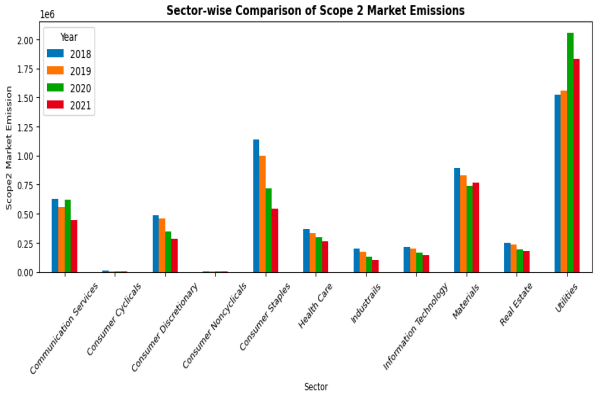
Further each financial variable chosen for the hypothesis testing was analysed for mean sector performance. For instance the results Return on Assets showed highest performance for the Healthcare and Information Technology sector.

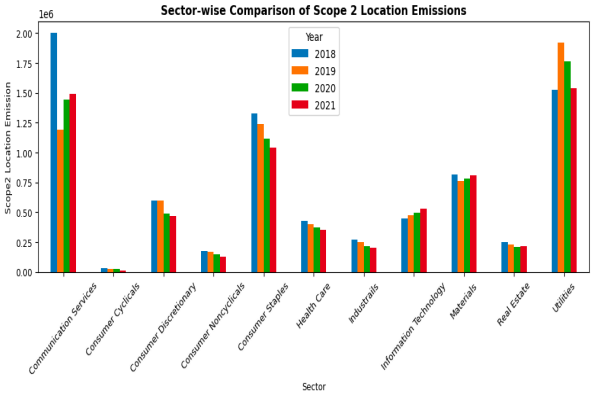


Time series of Scope 1 emissions showed decreasing trend from 2018 to 2020 but plateauing further until 2021. But Scope 2 emissions did not showed flat trend with no prominent decline in the emission levels.



Below plots show the Sector-wise comparison of Scope 2 Market and Scope 2 Location emission, where the Utilities sector comes out as the major contributor to the Scope 2 emissions for the period of analysis from 2018 to 2021.





Panel OLS Regression Results

The panel OLS regression technique was employed to estimate five different models (Model 1, Model 2, Model 3, Model 4, and Model 5) using a dataset consisting of 512 observations repeated without constant. The R-squared values of these models ranged from 0.1083 to 0.1787, suggesting that the independent variables collectively explain between 10.83% and 17.87% of the variation observed in the dependent variable. Furthermore, the F-statistics for all models yielded significant results, with p-values of 0, indicating that the regression models as a whole are statistically significant. These findings emphasize the relationships between the independent variables and the dependent variable in each respective model.

Model Comparison 1					
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Dep. Variable	**log(Scope1)**	**log(Scope2 Location)**	**log(Scope2 Market)**	**log(Sum Emission Location)**	**log(Sum Emission Market)**
Estimator	Panel OLS	Panel OLS	Panel OLS	Panel OLS	Panel OLS
No. Observations	512	512	512	512	512
Cov. Est.	Unadjusted	Unadjusted	Unadjusted	Unadjusted	Unadjusted
R-squared	0.1787	0.1083	0.111	0.1624	0.1649
R-Squared (Within)	-4.2813	-1.258	-0.1193	-2.8869	-1.5943
R-Squared (Between)	0.2161	0.1348	0.1329	0.1966	0.1996
R-Squared (Overall)	0.1787	0.1083	0.111	0.1624	0.1649
F-statistic	13.678	7.6332	7.8467	12.186	12.413
P-value (F-stat)	0	0	0	0	0
Constant	14.487***	13.455***	12.270***	14.871***	14.564***
	-36.23	-46.365	-31.718	-46.322	-41.964
Asset Turnover	-0.6201***	-0.5035***	-0.0957	-0.6127***	-0.4544***
	(-3.1440)	(-3.5180)	(-0.5013)	(-3.8696)	(-2.6544)
Current Ratio	-0.1906	0.0391	0.0197	0.0019	-0.0147
	(-1.2819)	-0.3625	-0.1367	-0.0162	(-0.1143)
EV to EBIT	-0.0048***	-0.0011	-0.0012	-0.0022**	-0.0027**
	(-3.4329)	(-1.1116)	(-0.8605)	(-2.0118)	(-2.2205)
Interest Coverage Ratio	-0.0005*	-0.0005**	-0.0001	-0.0006**	-0.0005*
	(-1.7226)	(-2.2680)	(-0.4232)	(-2.3110)	(-1.6654)
Price to Cash Flow per Share	-0.0561***	-0.0383***	-0.0407***	-0.0458***	-0.0480***
	(-4.7436)	(-4.4648)	(-3.5590)	(-4.8196)	(-4.6745)
Return on Assets	-0.0648***	-0.0194	-0.0938***	-0.0438**	-0.0704***
	(-3.0393)	(-1.2557)	(-4.5477)	(-2.5575)	(-3.8011)
Total Debt to Total Equity	0.0001	0.0006	0.0011	0.0007	0.0007
	-0.1323	-0.9985	-1.3771	-1.085	-1.0541
Weighted Average Cost of Capital	-0.0717	-0.0168	0.0909*	-0.0562	-0.0381
	(-1.4017)	(-0.4522)	-1.8382	(-1.3683)	(-0.8582)
T-stats reported in parentheses , ***p < 0.01, **p < 0.05, *p < 0.10					

Upon analysing the model comparison 1, it is evident that several variables exhibit statistical significance at a 1% significance level (p-value < 0.01). Specifically, Asset Turnover demonstrates statistical significance in both Model 1 and Model 2, suggesting that it is positively associated with the dependent variable (Scope1 emissions). Additionally, EV to EBIT is statistically significant in Model 1, indicating positive association with the Scope1 emissions. Price to Cash Flow per Share emerges as consistently significant across all models, indicating positive association with carbon emissions. Return on Assets also displays statistical significance in Model 1, indicating its relevance in predicting Scope1 emissions. On the other hand, variables such as Current Ratio, Interest Coverage Ratio, Total Debt to Total Equity, and Weighted Average Cost of Capital do not exhibit statistical significance at the 1% level, suggesting that they may have limited impact on Scope1 emissions in the given models.

Model Comparison 2					
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Dep. Variable	**log(Scope1)**	**log(Scope2 Location)**	**log(Scope2 Market)**	**log(Sum Emission Loc)**	**log(Sum Emission Mkt)**
Estimator	PanelOLS	PanelOLS	PanelOLS	PanelOLS	PanelOLS
No. Observations	512	512	512	512	512
Cov. Est.	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.8843	0.9156	0.9047	0.9084	0.9042
R-Squared (Within)	-43.927	-38.764	-3.9114	-59.504	-27.24
R-Squared (Between)	0.8989	0.9292	0.9197	0.9224	0.9181
R-Squared (Overall)	0.8843	0.9156	0.9047	0.9084	0.9042
F-statistic	481.71	683.55	598.26	625.03	594.79
P-value (F-stat)	0	0	0	0	0
Asset Turnover	2.0125***	1.9415***	2.1341***	2.0896***	2.1922***
	-2.8417	-2.9848	-3.2344	-2.9924	-3.0938
Current Ratio	1.9672***	2.0431***	1.8472***	2.2169***	2.1545***
	-4.5516	-6.2419	-5.4509	-5.8043	-5.4649
Enterprise Value to EBIT	0.0007	0.0040**	0.0035*	0.0034**	0.0028
	-0.4154	-2.4095	-1.8376	-1.9802	-1.5892
Interest Coverage Ratio	-0.0006	-0.0006**	-0.0002	-0.0006**	-0.0005
	(-1.6387)	(-2.0266)	(-0.6682)	(-2.0110)	(-1.6231)
Price to Cash Flow per Share	-0.0001	0.0137	0.0067	0.0117	0.0083

	(-0.0045)	-0.6038	-0.2802	-0.4717	-0.3326
Return on Assets	-0.0047	0.0364	-0.0429	0.0179	-0.0099
	(-0.0920)	-0.7857	(-0.7529)	-0.3682	(-0.1949)
Total Debt to Total Equity	0.0127*	0.0122*	0.0117**	0.0136*	0.0134*
	-1.9216	-1.9349	-1.9851	-1.9389	-1.9599
Weighted Average Cost of Capital	0.9177***	0.9021***	0.9289***	0.9594***	0.9566***
	-8.9104	-9.3089	-9.5588	-9.4285	-9.4408
T-stats reported in parentheses , ***p < 0.01, **p < 0.05, *p < 0.10					

In Model Comparison 2, a thorough assessment of the variable significance based on their p-values reveals key findings. At a 1% significance level, several variables display statistically significant relationships with the dependent variable. These include Asset Turnover and Current Ratio, which show positive relationship with carbon emissions. Additionally, the Weighted Average Cost of Capital showed positive association with carbon emission across all models. Moreover, the variables Enterprise Value to EBIT, Interest Coverage Ratio, and Total Debt to Total Equity display significance in specific models. However, the variables Price to Cash Flow per Share and Return on Assets do not reach the 1% significance level in any of the models, suggesting a weaker relationship with the dependent variable. These findings provide valuable insights into the variables that have significant positive association with the dependent variable in Model Comparison 2, facilitating a more nuanced understanding of the underlying relationships within the data

Conclusion

In conclusion, upon analyzing both Model Comparison 1 and Model Comparison 2, it is evident that certain variables consistently exhibit statistical significance at a 1% level. Across both models, Asset Turnover and Current Ratio consistently demonstrate significance, indicating their strong impact on Scope1 emissions. The Weighted Average Cost of Capital also consistently shows significance, highlighting its predictive power. In Model Comparison 1, EV to EBIT and Price to Cash Flow per Share are significant predictors, while Return on Assets is significant in Model Comparison 2. On the other hand, variables such as Interest Coverage Ratio, Total Debt to Total Equity, and certain others do not exhibit significant relationships with Scope1 emissions.

These findings provide valuable insights into the variables that have the greatest influence on Scope1 emissions, helping to inform decision-making and further analysis in the context of the given models.

When interpreting the results of the analysis, it is important to acknowledge the limitations of the data and consider the quality of the data used. One of the limitations is data availability, as the analysis relies on the presence of data for the variables included in the models. Additionally, the accuracy and reliability of the data can impact the validity of the analysis, as errors, inconsistencies, or missing values may introduce biases or affect the statistical results. The scope and representativeness of the data are also important considerations, as the findings may not be generalized to broader populations or different contexts. Furthermore, the timeframe of the data used may restrict the generalizability of the findings to other time periods. Regarding data quality, accuracy, completeness, consistency, and reliability are crucial factors. Ensuring accurate representation, minimal missing values, consistency in measurements, and reliable data sources contribute to the overall quality of the data. Being aware of these data limitations and considerations is essential in interpreting the results and understanding the implications of the analysis.

References

Oyewo, B. (2023). *Shibboleth Authentication Request*. [online] Oclc.org. Available at: https://www-sciencedirect-com.ucd.idm.oclc.org/science/article/pii/S0301479723002621?via%3Dihub [Accessed 14 Jun. 2023].

Ogunrinde, O., Shittu, E. and Dhanda, K.K. (2022). Distilling the Interplay Between Corporate Environmental Management, Financial, and Emissions Performance: Evidence From U.S. Firms. *IEEE Transactions on Engineering Management*, 69(6), pp.1–29. doi:https://doi.org/10.1109/tem.2020.3040158.

King, A.A. and Lenox, M.J. (2001). Does It Really Pay to Be Green? An Empirical Study of Firm Environmental and Financial Performance: An Empirical Study of Firm Environmental and

Financial Performance. *Journal of Industrial Ecology*, 5(1), pp.105–116. doi:https://doi.org/10.1162/108819801753358526.

Konadu, R., Ahinful, G.S., Boakye, D.J. and Elbardan, H. (2022). Board gender diversity, environmental innovation and corporate carbon emissions. *Technological Forecasting and Social Change*, 174(January 2022, 121279), p.121279. doi:https://doi.org/10.1016/j.techfore.2021.121279.

Appendix

Variable	Definition	Source
Scope 1	Natural logarithm of direct emissions from sources owned or controlled by the firm (in tonnes of carbon dioxide equivalent)	Race to Zero data explorer
Scope 2 Location	Natural logarithm of Location-based Scope 2 emissions are emissions calculated based on the average emissions intensity of a local power grid (in tonnes of carbon dioxide equivalent)	Race to Zero data explorer
Scope 2 Market	Natural logarithm of Market-based Scope 2 emissions are emissions calculated based on a specific purchase contract or agreement for energy (in tonnes of carbon dioxide equivalent)	Race to Zero data explorer
Asset Turnover	Asset turnover, total asset turnover, or asset turns is a financial ratio that measures the efficiency of a company's use of its assets	Refinitive
Current Ratio	The current ratio is a liquidity ratio that measures resources to meet its short-term obligations	Refinitive
EV to EBIT	The EV/EBIT ratio compares a company's enterprise value (EV) to its earnings before interest and taxes (EBIT).	Refinitive
Interest Coverage Ratio	Debt and profitability ratio, measures the ability to pay interest on its outstanding debt.	Refinitive
Price to Cash Flow per Share	Price-to-cash flow (P/CF) ratio measures the value of a stock's price relative to its operating cash flow per share	Refinitive
Return on Assets	Net profit/net assets, measures profitability of company is in relation to its total assets.	Refinitive
Total Debt to Total Equity	D/E ratio indicates debt of company compared to its assets	Refinitive
Weighted average cost of capital (WACC)	Average rate that a business pays to finance its assets	Refinitive

GDSProject

June 15, 2023

1 FIN41910 Green Data Science: Group Project

Submitted by: * Aman Parary (18417714) * Shraddha Nagar (22200477) * Rohan Kuntoji (22202093) * William Fernandes (22201553)

[1]:
```
import pandas as pd
import numpy as np
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.common.by import By
from fake_useragent import UserAgent
from time import sleep
import matplotlib.pyplot as plt
from matplotlib.colors import LinearSegmentedColormap
import seaborn as sns
import statsmodels.api as sm
from linearmodels.panel import PanelOLS
from linearmodels.panel.results import compare
```

1.1 Emission Data

1.1.1 Web-Scrape Carbon Emission Data

[2]:
```
#######################################################
##   Run only to web-scrape data from Race to Zero   ##
#######################################################

user_agent = UserAgent()
user_agent_value = user_agent.random
options = webdriver.ChromeOptions()
options.add_argument(f'user-agent={user_agent_value}')
driver = webdriver.Chrome(options=options)
driver.get("https://racetozerodataexplorer.org/proceed/")

#total_height = driver.execute_script("return document.body.scrollHeight")
total_height = 50000
```

```python
# Set the amount to scroll on each iteration
scroll_height = 200  # You can adjust this value

# Initialize current height to 0
current_height = 0

# Scroll the page slowly
while current_height < total_height:
    driver.execute_script(f"window.scrollBy(0, {scroll_height});")
    current_height += scroll_height

driver.execute_script("window.scrollTo(0, 0);")

html_content = driver.page_source
soup = BeautifulSoup(html_content, "html.parser")

# Find the table on the page
table = soup.find_all('table')[0]

# Find the table headers
headers = [header.text for header in table.find('thead').find_all('th')]

# Get rows in the table body
rows = table.find_all('tr')
# rows = rows[0:4]

data = []
for i, row in enumerate(rows[1:], start=0):
    cols = row.find_all('th')
    company_name = ' '.join(ele.text.strip() for ele in cols)

    # Initialize the data dictionary for this row
    row_data = {
        'Company': company_name,
        'Scope1 2018': '', 'Scope1 2019': '', 'Scope1 2020': '', 'Scope1 2021':␣
↪'',
        'Scope2 Location 2018': '', 'Scope2 Location 2019': '', 'Scope2␣
↪Location 2020': '', 'Scope2 Location 2021': '',
        'Scope2 Market 2018': '', 'Scope2 Market 2019': '', 'Scope2 Market␣
↪2020': '', 'Scope2 Market 2021': '',
        'Scope3': ''
    }

    # Scopes
    scopes = ['Scope1', 'Scope2 Location', 'Scope2 Market', 'Scope3']
```

```
    td_elements = row.find_all('td')
    for j, td in enumerate(td_elements[:4], start=0):  # replace with␣
↪td_elements to process all tds
        classes = td.get('class', [])
        if j == 3:
            if td.find(class_='styles_value__rPIE4'):
                value = td.find(class_='styles_value__rPIE4').text
                row_data[f'{scopes[j]}'] = value
        elif any('styles_hasButton' in class_name for class_name in classes):
            xpath = f"(//tr[{i+1}]//td)[{j+1}]//button"
            button = driver.find_element(By.XPATH, xpath)
            button.click()
            wait = WebDriverWait(driver, 20)
            element = wait.until(EC.visibility_of_element_located((By.ID,␣
↪'headlessui-portal-root')))

            # Extract data from the modal
            modal_html = driver.page_source
            modal_soup = BeautifulSoup(modal_html, 'html.parser')
            modal = modal_soup.find('div', id='headlessui-portal-root')


            # Get the data for the current scope
            data_elements = modal.find_all('div',␣
↪class_='styles_axisItem__tJkDc')

            for element in data_elements:
                year = element.find(class_='styles_axisDate__YlTNP').text
                value = element.find(class_='styles_axisValue__ipOq1').text
                row_data[f'{scopes[j]} {year}'] = value

            # sleep(3)
            close_button = driver.find_element(By.CSS_SELECTOR, 'button.
↪styles_close__cpVjT')
            close_button.click()

    data.append(row_data)
    # if (i > 0 and i % 20 == 0):
    #     sleep(10)
# Print data for debugging
# for row in data:
#     print(row)
#     print('\n\n\n')


df = pd.DataFrame(data)
```

```
df = df.mask(df == '')
df = df.mask(df == '-')
cols = df.columns.difference(['Company'])
# convert the relevant columns
for i in cols:
    df[i] = df[i].str.replace(',','')

df.to_excel('Emissions_Data.xlsx',index=False)
```

1.1.2 Emission Data: Cleaning and Organising

[3]:
```
df = pd.read_excel('Emissions_Data.xlsx')
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 14 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Company               500 non-null    object
 1   Scope1 2018           321 non-null    float64
 2   Scope1 2019           370 non-null    float64
 3   Scope1 2020           397 non-null    float64
 4   Scope1 2021           413 non-null    float64
 5   Scope2 Location 2018  340 non-null    float64
 6   Scope2 Location 2019  389 non-null    float64
 7   Scope2 Location 2020  419 non-null    float64
 8   Scope2 Location 2021  425 non-null    float64
 9   Scope2 Market 2018    264 non-null    float64
 10  Scope2 Market 2019    311 non-null    float64
 11  Scope2 Market 2020    340 non-null    float64
 12  Scope2 Market 2021    351 non-null    float64
 13  Scope3                422 non-null    float64
dtypes: float64(13), object(1)
memory usage: 54.8+ KB
```

[4]:
```
df[df.Company=='Interpublic Group of Companies, Inc.']

# Dropping Row Index 232
df = df.drop([df.index[232]]).reset_index(drop=True)
df
```

[4]:
```
                          Company  Scope1 2018  Scope1 2019  Scope1 2020  \
0    37 Interactive Entertainment          NaN        129.0        169.0
1            A.P. Moller - Maersk   38749000.0   36204000.0   33902000.0
2                          AAK AB          NaN          NaN          NaN
3                        AB Volvo          NaN     245000.0     205000.0
```

4	ABB	375000.0	257500.0	225547.0
..	…	…	…	…
494	Yum China Holdings, Inc.	NaN	114720.0	135615.0
495	Yum! Brands, Inc.	48047.0	23545.0	38907.0
496	Yutong Bus Co., Ltd.	NaN	NaN	NaN
497	Żabka Polska Sp. z o. o.	NaN	NaN	NaN
498	ZF Friedrichshafen AG	NaN	NaN	384000.0

	Scope1 2021	Scope2 Location 2018	Scope2 Location 2019 \
0	152.0	NaN	2701.0
1	36863000.0	338000.0	287000.0
2	269380.0	NaN	NaN
3	245000.0	NaN	215000.0
4	190900.0	650799.0	456000.0
..	…	…	…
494	158887.0	NaN	1956072.0
495	39013.0	155439.0	140341.0
496	NaN	NaN	NaN
497	7653.0	NaN	NaN
498	416000.0	NaN	NaN

	Scope2 Location 2020	Scope2 Location 2021	Scope2 Market 2018 \
0	2395.0	1794.0	NaN
1	289000.0	310170.0	NaN
2	NaN	18632.0	NaN
3	193000.0	204000.0	NaN
4	414000.0	370000.0	627903.0
..	…	…	…
494	1873757.0	2158856.0	NaN
495	128129.0	117621.0	159403.0
496	NaN	NaN	NaN
497	NaN	NaN	NaN
498	1323000.0	1377000.0	NaN

	Scope2 Market 2019	Scope2 Market 2020	Scope2 Market 2021	Scope3
0	NaN	NaN	NaN	13208.0
1	NaN	337000.0	330182.0	28950000.0
2	NaN	NaN	74351.0	6733176.0
3	124000.0	121000.0	114000.0	286000000.0
4	410800.0	335300.0	214000.0	124729000.0
..	…	…	…	…
494	NaN	NaN	NaN	8446963.0
495	144313.0	130953.0	112665.0	35134502.0
496	NaN	NaN	NaN	NaN
497	NaN	NaN	NaN	NaN
498	NaN	1130000.0	941000.0	91672227.0

```
[499 rows x 14 columns]
```

[5]:
```python
df1 = df.iloc[:,1:13]
df1.columns = pd.MultiIndex.from_product([['Scope1',
                                           'Scope2 Location',
                                           'Scope2 Market'],
                                        ['2018','2019','2020','2021']],
                                       names=['Scopes','Year'])
df1.insert(loc=0, column='Company', value=df['Company'].values)
df1['Scope3'] = df['Scope3']
df1.set_index('Company')
```

[5]:
```
Scopes                            Scope1                                      \
Year                                2018        2019        2020        2021
Company
37 Interactive Entertainment         NaN       129.0       169.0       152.0
A.P. Moller - Maersk          38749000.0  36204000.0  33902000.0  36863000.0
AAK AB                               NaN         NaN         NaN    269380.0
AB Volvo                             NaN    245000.0    205000.0    245000.0
ABB                             375000.0    257500.0    225547.0    190900.0
…                                     …           …           …           …
Yum China Holdings, Inc.             NaN    114720.0    135615.0    158887.0
Yum! Brands, Inc.               48047.0     23545.0     38907.0     39013.0
Yutong Bus Co., Ltd.                 NaN         NaN         NaN         NaN
Żabka Polska Sp. z o. o.             NaN         NaN         NaN      7653.0
ZF Friedrichshafen AG                NaN         NaN    384000.0    416000.0

Scopes                        Scope2 Location                                 \
Year                                     2018       2019       2020       2021
Company
37 Interactive Entertainment              NaN     2701.0     2395.0     1794.0
A.P. Moller - Maersk                 338000.0   287000.0   289000.0   310170.0
AAK AB                                    NaN        NaN        NaN    18632.0
AB Volvo                                  NaN   215000.0   193000.0   204000.0
ABB                                  650799.0   456000.0   414000.0   370000.0
…                                          …          …          …          …
Yum China Holdings, Inc.                  NaN  1956072.0  1873757.0  2158856.0
Yum! Brands, Inc.                    155439.0   140341.0   128129.0   117621.0
Yutong Bus Co., Ltd.                      NaN        NaN        NaN        NaN
Żabka Polska Sp. z o. o.                  NaN        NaN        NaN        NaN
ZF Friedrichshafen AG                     NaN        NaN  1323000.0  1377000.0

Scopes                        Scope2 Market                                   \
Year                                   2018       2019       2020       2021
Company
37 Interactive Entertainment            NaN        NaN        NaN        NaN
A.P. Moller - Maersk                    NaN        NaN   337000.0  330182.0
```

```
AAK AB                                  NaN       NaN        NaN   74351.0
AB Volvo                                NaN  124000.0   121000.0  114000.0
ABB                                627903.0  410800.0   335300.0  214000.0
…                                       …        …          …         …
Yum China Holdings, Inc.                NaN       NaN        NaN       NaN
Yum! Brands, Inc.                  159403.0  144313.0   130953.0  112665.0
Yutong Bus Co., Ltd.                    NaN       NaN        NaN       NaN
Żabka Polska Sp. z o. o.                NaN       NaN        NaN       NaN
ZF Friedrichshafen AG                   NaN       NaN  1130000.0  941000.0

Scopes                              Scope3
Year
Company
37 Interactive Entertainment       13208.0
A.P. Moller - Maersk            28950000.0
AAK AB                           6733176.0
AB Volvo                       286000000.0
ABB                            124729000.0
…                                        …
Yum China Holdings, Inc.         8446963.0
Yum! Brands, Inc.               35134502.0
Yutong Bus Co., Ltd.                   NaN
Żabka Polska Sp. z o. o.               NaN
ZF Friedrichshafen AG           91672227.0

[499 rows x 13 columns]
```

[6]:
```python
Scope1 = df1['Scope1']
Scope1.insert(loc=0, column='Company', value=df['Company'].values)
Scope1.set_index('Company',inplace=True)
Scope1 = pd.DataFrame(Scope1.stack(level=-1),columns = ['Scope1'])
Scope1
```

[6]:
```
                                      Scope1
Company                      Year
37 Interactive Entertainment 2019       129.0
                             2020       169.0
                             2021       152.0
A.P. Moller - Maersk         2018  38749000.0
                             2019  36204000.0
…                                          …
Yum! Brands, Inc.            2020     38907.0
                             2021     39013.0
Żabka Polska Sp. z o. o.     2021      7653.0
ZF Friedrichshafen AG        2020    384000.0
                             2021    416000.0
```

```
[1498 rows x 1 columns]
```

[7]:
```
Scope2Location = df1['Scope2 Location']
Scope2Location.insert(loc=0, column='Company', value=df['Company'].values)
Scope2Location.set_index('Company',inplace=True)
Scope2Location = pd.DataFrame(Scope2Location.stack(level=-1),columns=['Scope2␣
↪Location'])
Scope2Location
```

[7]:
```
                                    Scope2 Location
Company                       Year
37 Interactive Entertainment 2019           2701.0
                             2020           2395.0
                             2021           1794.0
A.P. Moller - Maersk         2018         338000.0
                             2019         287000.0
…                                              …
Yum! Brands, Inc.            2019         140341.0
                             2020         128129.0
                             2021         117621.0
ZF Friedrichshafen AG        2020        1323000.0
                             2021        1377000.0

[1570 rows x 1 columns]
```

[8]:
```
Scope2Market = df1['Scope2 Market']
Scope2Market.insert(loc=0, column='Company', value=df['Company'].values)
Scope2Market.set_index('Company',inplace=True)
Scope2Market = pd.DataFrame(Scope2Market.stack(level=-1),columns=['Scope2␣
↪Market'])
Scope2Market
```

[8]:
```
                             Scope2 Market
Company               Year
A.P. Moller - Maersk  2020       337000.0
                      2021       330182.0
AAK AB                2021        74351.0
AB Volvo              2019       124000.0
                      2020       121000.0
…                                     …
Yum! Brands, Inc.     2019       144313.0
                      2020       130953.0
                      2021       112665.0
ZF Friedrichshafen AG 2020      1130000.0
                      2021       941000.0

[1263 rows x 1 columns]
```

[9]:
```
company = list(df.Company)
sector_data = pd.read_excel('GDS sample data.xlsx')
sector_data.columns = ['Company','Sector']
sector_data = sector_data.sort_values('Company')
sector_data.drop_duplicates(keep='first', inplace=True)
sector_data.set_index('Company',inplace=True,drop=True)
sector_data['2018'] = sector_data['Sector']
sector_data['2019'] = sector_data['Sector']
sector_data['2020'] = sector_data['Sector']
sector_data['2021'] = sector_data['Sector']
sector_data.drop(['Sector'],axis=1,inplace=True)
# Create multi-index column header
header = pd.MultiIndex.from_tuples([('Year', '2018'), ('Year', '2019'),
                                    ('Year', '2020'), ('Year', '2021')])

# Assign the multi-index column header to the DataFrame
sector_data.columns = pd.MultiIndex.from_tuples([col for col in sector_data.
 ↪columns])
sector_data.columns = header

sector_data = pd.DataFrame(sector_data.stack(level=-1))
sector_data.reset_index(drop=False,inplace=True)
sector_data.rename(columns={'Year':'Sector','level_1':'Year'},inplace=True)
sector_data.set_index(['Company','Year'],inplace=True)
sector_data
```

[9]:
```
                                                   Sector
Company                       Year
37 Interactive Entertainment 2018  Communication Services
                             2019  Communication Services
                             2020  Communication Services
                             2021  Communication Services
A.P. Moller - Maersk         2018             Industrails
…                                                      …
Ørsted                       2021               Utilities
Żabka Polska Sp. z o. o.     2018   Consumer Noncyclicals
                             2019   Consumer Noncyclicals
                             2020   Consumer Noncyclicals
                             2021   Consumer Noncyclicals

[1996 rows x 1 columns]
```

[10]:
```
sector_data.shape
```

[10]: (1996, 1)

[11]:
```
data1 = pd.merge(sector_data, Scope1, how='left',on=['Company','Year'])
data2 = pd.merge(data1, Scope2Location, how='left',on=['Company','Year'])
emission_data = pd.merge(data2, Scope2Market, how='left',on=['Company','Year'])
```

[12]:
```
emission_data
```

[12]:
```
                                                  Sector      Scope1  \
Company                       Year
37 Interactive Entertainment 2018  Communication Services         NaN
                             2019  Communication Services       129.0
                             2020  Communication Services       169.0
                             2021  Communication Services       152.0
A.P. Moller - Maersk         2018             Industrails  38749000.0
…                                                      …         …
Ørsted                       2021               Utilities   2142000.0
Żabka Polska Sp. z o. o.     2018   Consumer Noncyclicals         NaN
                             2019   Consumer Noncyclicals         NaN
                             2020   Consumer Noncyclicals         NaN
                             2021   Consumer Noncyclicals      7653.0

                                   Scope2 Location  Scope2 Market
Company                       Year
37 Interactive Entertainment 2018              NaN            NaN
                             2019           2701.0            NaN
                             2020           2395.0            NaN
                             2021           1794.0            NaN
A.P. Moller - Maersk         2018         338000.0            NaN
…                                                …            …
Ørsted                       2021          53000.0          802.0
Żabka Polska Sp. z o. o.     2018              NaN            NaN
                             2019              NaN            NaN
                             2020              NaN            NaN
                             2021              NaN            NaN

[1996 rows x 4 columns]
```

[13]:
```
##### Checking for missing values
df2 = df1.drop('Scope3',axis=1).set_index('Company')
drop_list = df2[df2.isna().any(axis=1)]
drop_list.reset_index(drop=False,inplace=True)
drop_list = list(drop_list['Company'])
##### List of Companies with available Emissions data
companies_list = emission_data.drop(drop_list, level=0, axis=0)
companies_list
```

```
/var/folders/9p/xr85g4y15w1fy3_389l3s_nh0000gn/T/ipykernel_86452/2857731499.py:2
: PerformanceWarning: dropping on a non-lexsorted multi-index without a level
```

```
parameter may impact performance.
  df2 = df1.drop('Scope3',axis=1).set_index('Company')
```

[13]:

Company	Year	Sector	Scope1	Scope2 Location	Scope2 Market
ABB	2018	Industrails	375000.0	650799.0	627903.0
	2019	Industrails	257500.0	456000.0	410800.0
	2020	Industrails	225547.0	414000.0	335300.0
	2021	Industrails	190900.0	370000.0	214000.0
AENA SME SA	2018	Industrails	30707.0	236695.0	218937.0
…		…	…	…	…
thyssenkrupp AG	2021	Materials	23200000.0	1300000.0	1500000.0
Ørsted	2018	Utilities	3483000.0	151000.0	46000.0
	2019	Utilities	1846000.0	123000.0	4000.0
	2020	Utilities	1851000.0	111000.0	1790.0
	2021	Utilities	2142000.0	53000.0	802.0

[856 rows x 4 columns]

[14]:

```
Tickers = pd.read_excel("Company_Financial_v3.xlsx", sheet_name=0)
Tickers.columns = ['Company','2018']
Tickers['2019'] = Tickers['2018']
Tickers['2020'] = Tickers['2018']
Tickers['2021'] = Tickers['2018']
Tickers = Tickers.sort_values('Company')
Tickers.set_index('Company',inplace=True)
header = pd.MultiIndex.from_tuples([('Year', '2018'), ('Year', '2019'),
                                    ('Year', '2020'), ('Year', '2021')])
# Assign the multi-index column header to the DataFrame
Tickers.columns = pd.MultiIndex.from_tuples([col for col in Tickers.columns])
Tickers.columns = header
Tickers = pd.DataFrame(Tickers.stack(level=-1))
Tickers.reset_index(drop=False,inplace=True)
Tickers.rename(columns={'Year':'Ticker','level_1':'Year'},inplace=True)
#Tickers['Year'] = pd.to_datetime(Tickers['Year']).dt.year
Tickers.set_index(['Company','Year'],inplace=True)
Tickers
```

[14]:

Company	Year	Ticker
ABB	2018	ABBNY.PK
	2019	ABBNY.PK
	2020	ABBNY.PK
	2021	ABBNY.PK
AENA SME SA	2018	AENA.MC
…		…
thyssenkrupp AG	2021	TKAG.DE

```
Ørsted          2018  ORSTED.CO
                2019  ORSTED.CO
                2020  ORSTED.CO
                2021  ORSTED.CO

[840 rows x 1 columns]
```

[15]:
```python
data = pd.merge(companies_list, Tickers, on=['Company','Year'],how='left')
data.reset_index(inplace=True, drop=False)
data.sort_values('Ticker')
data['Year'] = pd.to_datetime(data['Year']).dt.year
data.set_index(['Ticker','Sector','Year'],inplace=True,drop=True)
data.drop(['Company'],axis=1,inplace=True)
data['Sum_Emission_Mkt'] = data['Scope1'] + data['Scope2 Market']
data['Sum_Emission_Loc'] = data['Scope1'] + data['Scope2 Location']
data
```

[15]:
```
                                Scope1  Scope2 Location  Scope2 Market  \
Ticker    Sector      Year
ABBNY.PK  Industrails 2018    375000.0         650799.0       627903.0
                      2019    257500.0         456000.0       410800.0
                      2020    225547.0         414000.0       335300.0
                      2021    190900.0         370000.0       214000.0
AENA.MC   Industrails 2018     30707.0         236695.0       218937.0
…                                  …               …            …
TKAG.DE   Materials   2021  23200000.0        1300000.0      1500000.0
ORSTED.CO Utilities   2018   3483000.0         151000.0        46000.0
                      2019   1846000.0         123000.0         4000.0
                      2020   1851000.0         111000.0         1790.0
                      2021   2142000.0          53000.0          802.0

                             Sum_Emission_Mkt  Sum_Emission_Loc
Ticker    Sector      Year
ABBNY.PK  Industrails 2018         1002903.0         1025799.0
                      2019          668300.0          713500.0
                      2020          560847.0          639547.0
                      2021          404900.0          560900.0
AENA.MC   Industrails 2018          249644.0          267402.0
…                                       …               …
TKAG.DE   Materials   2021        24700000.0        24500000.0
ORSTED.CO Utilities   2018         3529000.0         3634000.0
                      2019         1850000.0         1969000.0
                      2020         1852790.0         1962000.0
                      2021         2142802.0         2195000.0

[856 rows x 5 columns]
```

1.2 Financial Data

[16]:
```
Fin_df = pd.read_excel("Company_Financial_v3.xlsx", sheet_name=2)
Fin_df.head()
```

[16]:
```
    Ticker  Currency       Asset Turnover     Asset Turnover.1  \
0   Ticker  Currency  2016-06-30 00:00:00  2017-06-30 00:00:00
1   MSFT.O       USD             0.495483              0.43522
2    PEP.O       USD                  NaN                  NaN
3    FGP.L       GBP                  NaN                  NaN
4  ENEI.MI       EUR                  NaN                  NaN

      Asset Turnover.2     Asset Turnover.3     Asset Turnover.4  \
0  2018-06-30 00:00:00  2019-06-30 00:00:00  2020-06-30 00:00:00
1             0.433498             0.461467             0.486556
2                  NaN                  NaN                  NaN
3                  NaN                  NaN                  NaN
4                  NaN                  NaN                  NaN

      Asset Turnover.5     Asset Turnover.6     Asset Turnover.7  …  \
0  2021-06-30 00:00:00  2018-12-31 00:00:00  2019-12-31 00:00:00  …
1             0.529336                  NaN                  NaN  …
2                  NaN                  NaN                  NaN  …
3                  NaN                  NaN                  NaN  …
4                  NaN             0.338598             0.323147  …

  Total Debt to Total Equity, Percent.115  \
0                     2021-02-27 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.116  \
0                     2017-07-31 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.117  \
0                     2018-07-31 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN
```

```
  Total Debt to Total Equity, Percent.118  \
0                     2019-07-31 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.119  \
0                     2020-07-31 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.120  \
0                     2016-11-30 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.121  \
0                     2017-11-30 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.122  \
0                     2019-11-30 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.123  \
0                     2020-11-30 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
4                                     NaN

  Total Debt to Total Equity, Percent.124
0                     2021-01-02 00:00:00
1                                     NaN
2                                     NaN
3                                     NaN
```

```
4                                       NaN

[5 rows x 1809 columns]
```

[17]:
```python
# Extract years and append to column names
for col in Fin_df.columns[2:]:
    year = pd.to_datetime(Fin_df.loc[0, col]).year
    Fin_df[col] = Fin_df[col].where(Fin_df.index != 0, Fin_df[col].name + ' ' +␣
 ↪str(year))

Fin_df.columns = Fin_df.iloc[0]
Fin_df = Fin_df.iloc[1:].reset_index(drop=True)
Fin_df.head()
```

[17]:
```
0    Ticker Currency Asset Turnover 2016 Asset Turnover.1 2017  \
0    MSFT.O      USD            0.495483               0.43522
1     PEP.O      USD                 NaN                   NaN
2     FGP.L      GBP                 NaN                   NaN
3   ENEI.MI      EUR                 NaN                   NaN
4  SAABb.ST      SEK                 NaN                   NaN

0 Asset Turnover.2 2018 Asset Turnover.3 2019 Asset Turnover.4 2020  \
0              0.433498              0.461467              0.486556
1                   NaN                   NaN                   NaN
2                   NaN                   NaN                   NaN
3                   NaN                   NaN                   NaN
4                   NaN                   NaN                   NaN

0 Asset Turnover.5 2021 Asset Turnover.6 2018 Asset Turnover.7 2019  …  \
0              0.529336                   NaN                   NaN  …
1                   NaN                   NaN                   NaN  …
2                   NaN                   NaN                   NaN  …
3                   NaN              0.338598              0.323147  …
4                   NaN              0.655736              0.610988  …

0 Total Debt to Total Equity, Percent.115 2021  \
0                                          NaN
1                                          NaN
2                                          NaN
3                                          NaN
4                                          NaN

0 Total Debt to Total Equity, Percent.116 2017  \
0                                          NaN
1                                          NaN
2                                          NaN
3                                          NaN
```

```
4                                           NaN

0 Total Debt to Total Equity, Percent.117 2018  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.118 2019  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.119 2020  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.120 2016  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.121 2017  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.122 2019  \
0                                           NaN
1                                           NaN
2                                           NaN
3                                           NaN
4                                           NaN

0 Total Debt to Total Equity, Percent.123 2020  \
0                                           NaN
1                                           NaN
```

```
2                                          NaN
3                                          NaN
4                                          NaN

0 Total Debt to Total Equity, Percent.124 2021
0                                          NaN
1                                          NaN
2                                          NaN
3                                          NaN
4                                          NaN

[5 rows x 1809 columns]
```

[18]:
```python
# Remove trailing digits and periods from metric names
Fin_df.columns = Fin_df.columns.str.replace('\.\d+', '')

# Melt DataFrame
df_melt = Fin_df.melt(id_vars=['Ticker', 'Currency'], var_name='Variable',␣
 ↪value_name='Value')

# Split variable into year and metric
split_variable = df_melt['Variable'].str.rsplit(' ', n=1, expand=True)
df_melt['Year'] = pd.to_datetime(split_variable[1]).dt.year
df_melt['Metric'] = split_variable[0]

# Drop the original variable column
df_melt.drop(columns='Variable', inplace=True)
df_melt.head()
```

```
/var/folders/9p/xr85g4y15w1fy3_389l3s_nh0000gn/T/ipykernel_86452/1132578631.py:2
: FutureWarning: The default value of regex will change from True to False in a
future version.
  Fin_df.columns = Fin_df.columns.str.replace('\.\d+', '')
```

[18]:

	Ticker	Currency	Value	Year	Metric
0	MSFT.O	USD	0.495483	2016	Asset Turnover
1	PEP.O	USD	NaN	2016	Asset Turnover
2	FGP.L	GBP	NaN	2016	Asset Turnover
3	ENEI.MI	EUR	NaN	2016	Asset Turnover
4	SAABb.ST	SEK	NaN	2016	Asset Turnover

[19]:
```python
company_df = pd.read_excel("Company_Financial_v2.xlsx", sheet_name=0)
# Extract only the relevant tickers we require
company_df = company_df[company_df['Ticker'].isin(Fin_df['Ticker'])].
 ↪reset_index(drop=True)
sector_df = pd.read_excel("GDS sample data.xlsx")
# Extract the relevant company name
```

```
sector_df = sector_df[sector_df['Name'].isin(company_df['Name'])].
 ↪reset_index(drop=True)
# Merge the relevant company name and its sector
comany_info_df = pd.merge(company_df, sector_df, on=['Name'])
comany_info_df = comany_info_df.drop_duplicates()
# Merge the company name and sector with their financials
df_merged = pd.merge(df_melt, comany_info_df, on='Ticker')
```

[20]:
```
df_merged
```

[20]:
```
         Ticker Currency     Value  Year                                Metric  \
0        MSFT.O      USD  0.495483  2016                        Asset Turnover
1        MSFT.O      USD   0.43522  2017                        Asset Turnover
2        MSFT.O      USD  0.433498  2018                        Asset Turnover
3        MSFT.O      USD  0.461467  2019                        Asset Turnover
4        MSFT.O      USD  0.486556  2020                        Asset Turnover
…          …       …        …    …                                   …
343325  CMPC.SN      USD       NaN  2016  Total Debt to Total Equity, Percent
343326  CMPC.SN      USD       NaN  2017  Total Debt to Total Equity, Percent
343327  CMPC.SN      USD       NaN  2019  Total Debt to Total Equity, Percent
343328  CMPC.SN      USD       NaN  2020  Total Debt to Total Equity, Percent
343329  CMPC.SN      USD       NaN  2021  Total Debt to Total Equity, Percent

                         Name                  Sector
0       Microsoft Corporation  Information Technology
1       Microsoft Corporation  Information Technology
2       Microsoft Corporation  Information Technology
3       Microsoft Corporation  Information Technology
4       Microsoft Corporation  Information Technology
…                         …                     …
343325          Empresas CMPC               Materials
343326          Empresas CMPC               Materials
343327          Empresas CMPC               Materials
343328          Empresas CMPC               Materials
343329          Empresas CMPC               Materials

[343330 rows x 7 columns]
```

[21]:
```
financial_df = df_merged.groupby(['Ticker', 'Name', 'Sector', 'Currency',␣
 ↪'Metric', 'Year'])['Value'].first().unstack('Metric')
financial_df
```

[21]:
```
Metric                                                          Asset
Turnover  \
Ticker Name                         Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
```

```
                                                                    2016
1.512548
                                                                    2017
1.47987
                                                                    2018
1.44318
                                                                    2019
1.399814
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
1.092159
                                                                    2018
1.204957
                                                                    2019
1.195812
                                                                    2020
1.01994
                                                                    2021
None

Metric                                                                   Current
Ratio  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
                                                                    2016
1.56892
                                                                    2017
1.5876
                                                                    2018
1.58839
                                                                    2019
1.63153
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
1.65807
                                                                    2018
0.92775
                                                                    2019
0.99091
                                                                    2020
1.00836
                                                                    2021
None
```

```
Metric
EV/Market Cap  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
                                                                   2016
1.101865
                                                                   2017
1.222738
                                                                   2018
1.399741
                                                                   2019
1.703784
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
1.305671
                                                                   2018
1.347702
                                                                   2019
1.329495
                                                                   2020
1.306896
                                                                   2021
None

Metric
Enterprise Value To EBIT (Daily Time Series Ratio)  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                   2016
10.437924
                                                                   2017
8.83933
                                                                   2018
8.873924
                                                                   2019
9.115623
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
32.133204
                                                                   2018
21.949991
                                                                   2019
21.706238
```

```
                                                                    2020
24.219634
                                                                    2021
24.284847

Metric                                                                   Interest
Coverage Ratio  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
                                                                    2016
24.347087
                                                                    2017
34.608333
                                                                    2018
38.227109
                                                                    2019
24.440199
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
3.566596
                                                                    2018
3.580645
                                                                    2019
3.699422
                                                                    2020
3.111111
                                                                    2021
None

Metric                                                                   Price To
Book Value Per Share (Daily Time Series Ratio)  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                    2016
1.172949
                                                                    2017
1.250662
                                                                    2018
0.832287
                                                                    2019
0.931975
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
```

```
-4.519815
                                                                  2018
-3.857731
                                                                  2019
-3.769444
                                                                  2020
-4.140058
                                                                  2021
-5.197306

Metric                                                                 Price To
Cash Flow Per Share (Daily Time Series Ratio)  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                  2016
6.008775
                                                                  2017
9.126765
                                                                  2018
18.462087
                                                                  2019
7.253062
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
72.373826
                                                                  2018
25.90865
                                                                  2019
24.122735
                                                                  2020
25.485675
                                                                  2021
23.339715

Metric                                                                 Return
On Assets - Actual  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                  2016
None
                                                                  2017
None
                                                                  2018
None
```

```
                                                                    2019
None
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
14.48
                                                                    2018
19.5
                                                                    2019
22.07
                                                                    2020
23.74
                                                                    2021
20.1

Metric                                                                   Total
Debt to Total Equity, Percent  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                    2016
None
                                                                    2017
59.571698
                                                                    2018
64.352752
                                                                    2019
78.153899
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
None
                                                                    2018
None
                                                                    2019
None
                                                                    2020
None
                                                                    2021
None

Metric                                                                   Weighted
Average Cost of Capital, (%)
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                    2016
```

```
4.052073
                                                                    2017
4.507422
                                                                    2018
5.749109
                                                                    2019
4.043643
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
6.126988
                                                                    2018
6.525537
                                                                    2019
4.150549
                                                                    2020
6.213764
                                                                    2021
2.038374

[1330 rows x 10 columns]
```

[22]:
```
# Drop these columns
financial_df = financial_df.drop(['EV/Market Cap', 'Price To Book Value Per␣
↪Share (Daily Time Series Ratio)'], axis=1)
```

[23]:
```
def convert_to_usd(row, metric):
    """
    Converts a value to USD given the currency and the metric

    Parameters:
        row (pandas.Series): A row of a DataFrame containing the data to be␣
↪converted
        metric (str): The name of the column representing the value to be␣
↪converted

    Returns:
        float: The converted value in USD

    """

    # Get the currency and the year of the row df
    currency = row.name[3]
    year = row.name[4]

    # If the currency is already USD or the metric is NaN, no need to convert
    if currency == 'USD' or pd.isna(row[metric]):
```

```
        return row[metric]

    # Convert the metric to USD using the hardcoded exchange rates
    try:
        exchange_rate = exchange_rates[year][currency]
    except KeyError:
        print(f'No exchange rate available for {currency} in year {year}')
        return np.nan

    return row[metric] * exchange_rate

def find_missing_values(df, metrics, num_years):
    """
    Finds and prints the tickers with missing values for specific metrics in a␣
↪DataFrame

    Parameters:
        df (pandas.DataFrame): The DataFrame containing the data.
        metrics (list): A list of strings representing the metrics to check for␣
↪missing values
        num_years (int): The minimum number of missing years to consider as␣
↪significant

    Returns:
        None
    """

    for metric in metrics:
        missing_values_dict = {}

        # Loop through each ticker
        for ticker in df.index.get_level_values('Ticker').unique():
            # Get the data for the current ticker
            ticker_data = df.loc[ticker]

            # Find the years where the metric is missing for the current ticker
            missing_years = ticker_data[ticker_data[f'{metric}'].isnull()].
↪index.tolist()

            # If there are more than num_years missing years, store them in the␣
↪dictionary
            if len(missing_years) > num_years:
                missing_values_dict[ticker] = missing_years

        # Print the tickers with the years they have missing values for the␣
↪given metric
```

```
        for ticker, years in missing_values_dict.items():
            print(f"Ticker {ticker} has missing values in {metric} for the␣
↪years {', '.join(map(str, years))}")

def count_no_missing_values(df, metrics):
    """
    Counts the number of tickers with no missing values for specific metrics in␣
↪a DataFrame

    Parameters:
        df (pandas.DataFrame): The DataFrame containing the data
        metrics (list): A list of strings representing the metrics to check for␣
↪missing values

    Returns:
        None

    """
    for metric in metrics:
        no_missing_values_tickers = []

        # Loop through each ticker
        for ticker in df.index.get_level_values('Ticker').unique():
            # Get the data for the current ticker
            ticker_data = df.loc[ticker]

            # Filter for the years 2017-2020
            ticker_data = ticker_data.loc[(ticker_data.index.
↪get_level_values('Year') >= 2017) & (ticker_data.index.
↪get_level_values('Year') <= 2021)]

            # Check if there are any missing values for the metric in these␣
↪years
            if ticker_data[f'{metric}'].isnull().sum() == 0:
                no_missing_values_tickers.append(ticker)

        # Print the number of tickers with no missing values for the given␣
↪metric in the years 2016-2020
        print(f"{len(no_missing_values_tickers)} tickers have no missing values␣
↪in {metric} for the years 2017-2021.")
```

[24]:
```
#####################################################
# USE THIS PART OF THE CODE TO CONVERT DATA TO USD #
#####################################################
```

```
exchange_rates = {
    2015: {'GBP': 1.5285, 'EUR': 1.1100, 'SEK': 0.1186, 'NOK': 0.1242, 'JPY': 0.
 ↪0083, 'CHF': 1.0403, 'TWD': 0.0315, 'DKK': 0.1488, 'MXN': 0.0631, 'CAD': 0.
 ↪7832, 'BRL': 0.305},
    2016: {'GBP': 1.3552, 'EUR': 1.1069, 'SEK': 0.1170, 'NOK': 0.1191, 'JPY': 0.
 ↪0092, 'CHF': 1.0153, 'TWD': 0.0310, 'DKK': 0.1487, 'MXN': 0.0536, 'CAD': 0.
 ↪7553, 'BRL': 0.2885},
    2017: {'GBP': 1.2890, 'EUR': 1.1301, 'SEK': 0.1173, 'NOK': 0.1211, 'JPY': 0.
 ↪0089, 'CHF': 1.0163, 'TWD': 0.0329, 'DKK': 0.1520, 'MXN': 0.0530, 'CAD': 0.
 ↪7713, 'BRL': 0.3134},
    2018: {'GBP': 1.3349, 'EUR': 1.1811, 'SEK': 0.1152, 'NOK': 0.1230, 'JPY': 0.
 ↪0091, 'CHF': 1.0225, 'TWD': 0.0332, 'DKK': 0.1584, 'MXN': 0.0521, 'CAD': 0.
 ↪7717, 'BRL': 0.2755},
    2019: {'GBP': 1.2772, 'EUR': 1.1199, 'SEK': 0.1058, 'NOK': 0.1137, 'JPY': 0.
 ↪0092, 'CHF': 1.0067, 'TWD': 0.0324, 'DKK': 0.1500, 'MXN': 0.0520, 'CAD': 0.
 ↪7538, 'BRL': 0.2538},
    2020: {'GBP': 1.2837, 'EUR': 1.1199, 'SEK': 0.1090, 'NOK': 0.1066, 'JPY': 0.
 ↪0094, 'CHF': 1.0665, 'TWD': 0.0340, 'DKK': 0.1532, 'MXN': 0.0469, 'CAD': 0.
 ↪7462, 'BRL': 0.1958},
    2021: {'GBP': 1.3757, 'EUR': 1.1830, 'SEK': 0.1166, 'NOK': 0.1164, 'JPY': 0.
 ↪0091, 'CHF': 1.0941, 'TWD': 0.0358, 'DKK': 0.1591, 'MXN': 0.0493, 'CAD': 0.
 ↪7978, 'BRL': 0.1855}
}

df_convert = financial_df.copy()

# Apply the function to each row of the DataFrame
metrics = financial_df.columns

# Commented the code below because we do not have any actual variables to be␣
 ↪converted to USD.
# We're using Ratios which are already unscaled.

# for metric in metrics:
#     # Convert metric to USD
#     df_convert[f'{metric} USD'] = financial_df.apply(convert_to_usd,␣
 ↪metric=metric, axis=1)
#     # Drop the original metric column
#     df_convert.drop(metric, axis=1, inplace=True)

df_convert
```

[24]:
```
Metric                                                          Asset
Turnover  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
```

```
None
                                                                    2016
1.512548
                                                                    2017
1.47987
                                                                    2018
1.44318
                                                                    2019
1.399814
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
1.092159
                                                                    2018
1.204957
                                                                    2019
1.195812
                                                                    2020
1.01994
                                                                    2021
None

Metric                                                                   Current
Ratio  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
                                                                    2016
1.56892
                                                                    2017
1.5876
                                                                    2018
1.58839
                                                                    2019
1.63153
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
1.65807
                                                                    2018
0.92775
                                                                    2019
0.99091
                                                                    2020
1.00836
                                                                    2021
None
```

```
Metric
Enterprise Value To EBIT (Daily Time Series Ratio)  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                   2016
10.437924
                                                                   2017
8.83933
                                                                   2018
8.873924
                                                                   2019
9.115623
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
32.133204
                                                                   2018
21.949991
                                                                   2019
21.706238
                                                                   2020
24.219634
                                                                   2021
24.284847

Metric                                                                 Interest
Coverage Ratio  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
None
                                                                   2016
24.347087
                                                                   2017
34.608333
                                                                   2018
38.227109
                                                                   2019
24.440199
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
3.566596
                                                                   2018
3.580645
                                                                   2019
```

3.699422
2020
3.111111
2021
None

Metric Price To
Cash Flow Per Share (Daily Time Series Ratio) \
Ticker Name Sector Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY 2015
NaN
2016
6.008775
2017
9.126765
2018
18.462087
2019
7.253062
…
…
YUM Yum! Brands, Inc. Consumer Discretionary USD 2017
72.373826
2018
25.90865
2019
24.122735
2020
25.485675
2021
23.339715

Metric Return
On Assets - Actual \
Ticker Name Sector Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY 2015
NaN
2016
None
2017
None
2018
None
2019
None
…
…

```
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
14.48
                                                                   2018
19.5
                                                                   2019
22.07
                                                                   2020
23.74
                                                                   2021
20.1

Metric                                                                 Total
Debt to Total Equity, Percent  \
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                   2016
None
                                                                   2017
59.571698
                                                                   2018
64.352752
                                                                   2019
78.153899
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
None
                                                                   2018
None
                                                                   2019
None
                                                                   2020
None
                                                                   2021
None

Metric                                                                 Weighted
Average Cost of Capital, (%)
Ticker Name                        Sector                 Currency Year
1911.T Sumitomo Forestry Co., Ltd. Consumer Discretionary JPY      2015
NaN
                                                                   2016
4.052073
                                                                   2017
4.507422
                                                                   2018
```

```
5.749109
                                                                    2019
4.043643
…
…
YUM    Yum! Brands, Inc.           Consumer Discretionary USD      2017
6.126988
                                                                    2018
6.525537
                                                                    2019
4.150549
                                                                    2020
6.213764
                                                                    2021
2.038374

[1330 rows x 8 columns]
```

[25]:
```
num_years = 2
find_missing_values(df_convert, metrics, num_years)
```

```
Ticker 1911.T has missing values in Asset Turnover for the years ('Sumitomo
Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2015), ('Sumitomo Forestry
Co., Ltd.', 'Consumer Discretionary', 'JPY', 2020), ('Sumitomo Forestry Co.,
Ltd.', 'Consumer Discretionary', 'JPY', 2021)
Ticker 6723.T has missing values in Asset Turnover for the years ('Renesas
Electronics Corporation', 'Information Technology', 'JPY', 2015), ('Renesas
Electronics Corporation', 'Information Technology', 'JPY', 2016), ('Renesas
Electronics Corporation', 'Information Technology', 'JPY', 2021)
Ticker UU.L has missing values in Asset Turnover for the years ('United
Utilities', 'Utilities', 'GBP', 2015), ('United Utilities', 'Utilities', 'GBP',
2016), ('United Utilities', 'Utilities', 'GBP', 2017), ('United Utilities',
'Utilities', 'GBP', 2018), ('United Utilities', 'Utilities', 'GBP', 2019)
Ticker 7731.T has missing values in Enterprise Value To EBIT (Daily Time Series
Ratio) for the years ('Nikon Corporation', 'Consumer Discretionary', 'JPY',
2015), ('Nikon Corporation', 'Consumer Discretionary', 'JPY', 2017), ('Nikon
Corporation', 'Consumer Discretionary', 'JPY', 2020)
Ticker AAL.O has missing values in Enterprise Value To EBIT (Daily Time Series
Ratio) for the years ('American Airlines Group Inc', 'Industrails', 'USD',
2015), ('American Airlines Group Inc', 'Industrails', 'USD', 2020), ('American
Airlines Group Inc', 'Industrails', 'USD', 2021)
Ticker ADSK.O has missing values in Enterprise Value To EBIT (Daily Time Series
Ratio) for the years ('Autodesk, Inc.', 'Information Technology', 'USD', 2015),
('Autodesk, Inc.', 'Information Technology', 'USD', 2016), ('Autodesk, Inc.',
'Information Technology', 'USD', 2017), ('Autodesk, Inc.', 'Information
Technology', 'USD', 2018)
Ticker NOKIA.HE has missing values in Enterprise Value To EBIT (Daily Time
```

Series Ratio) for the years ('Nokia Group', 'Information Technology', 'EUR', 2015), ('Nokia Group', 'Information Technology', 'EUR', 2016), ('Nokia Group', 'Information Technology', 'EUR', 2018)
Ticker TKAG.DE has missing values in Enterprise Value To EBIT (Daily Time Series Ratio) for the years ('thyssenkrupp AG', 'Materials', 'EUR', 2015), ('thyssenkrupp AG', 'Materials', 'EUR', 2019), ('thyssenkrupp AG', 'Materials', 'EUR', 2020)
Ticker UAL.O has missing values in Enterprise Value To EBIT (Daily Time Series Ratio) for the years ('United Airlines Holdings', 'Industrails', 'USD', 2015), ('United Airlines Holdings', 'Industrails', 'USD', 2020), ('United Airlines Holdings', 'Industrails', 'USD', 2021)
Ticker 1911.T has missing values in Interest Coverage Ratio for the years ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2015), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2020), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2021)
Ticker 6723.T has missing values in Interest Coverage Ratio for the years ('Renesas Electronics Corporation', 'Information Technology', 'JPY', 2015), ('Renesas Electronics Corporation', 'Information Technology', 'JPY', 2016), ('Renesas Electronics Corporation', 'Information Technology', 'JPY', 2021)
Ticker ADSK.O has missing values in Interest Coverage Ratio for the years ('Autodesk, Inc.', 'Information Technology', 'USD', 2015), ('Autodesk, Inc.', 'Information Technology', 'USD', 2016), ('Autodesk, Inc.', 'Information Technology', 'USD', 2017), ('Autodesk, Inc.', 'Information Technology', 'USD', 2018), ('Autodesk, Inc.', 'Information Technology', 'USD', 2019), ('Autodesk, Inc.', 'Information Technology', 'USD', 2020), ('Autodesk, Inc.', 'Information Technology', 'USD', 2021)
Ticker ETN has missing values in Interest Coverage Ratio for the years ('Eaton Corporation', 'Industrails', 'USD', 2015), ('Eaton Corporation', 'Industrails', 'USD', 2016), ('Eaton Corporation', 'Industrails', 'USD', 2017), ('Eaton Corporation', 'Industrails', 'USD', 2018), ('Eaton Corporation', 'Industrails', 'USD', 2019), ('Eaton Corporation', 'Industrails', 'USD', 2020), ('Eaton Corporation', 'Industrails', 'USD', 2021)
Ticker FGP.L has missing values in Interest Coverage Ratio for the years ('FirstGroup PLC', 'Industrails', 'GBP', 2015), ('FirstGroup PLC', 'Industrails', 'GBP', 2020), ('FirstGroup PLC', 'Industrails', 'GBP', 2021)
Ticker HPE has missing values in Interest Coverage Ratio for the years ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2015), ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2016), ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2017), ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2018), ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2019), ('Hewlett Packard Enterprise Company', 'Information Technology', 'USD', 2020)
Ticker LULU.O has missing values in Interest Coverage Ratio for the years ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2015), ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2016), ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2017), ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2018), ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2019), ('Lululemon Athletica Inc.', 'Consumer

Discretionary', 'USD', 2020), ('Lululemon Athletica Inc.', 'Consumer Discretionary', 'USD', 2021)
Ticker TGT has missing values in Interest Coverage Ratio for the years ('Target Corporation', 'Consumer Discretionary', 'USD', 2015), ('Target Corporation', 'Consumer Discretionary', 'USD', 2017), ('Target Corporation', 'Consumer Discretionary', 'USD', 2018)
Ticker AAL.O has missing values in Price To Cash Flow Per Share (Daily Time Series Ratio) for the years ('American Airlines Group Inc', 'Industrails', 'USD', 2015), ('American Airlines Group Inc', 'Industrails', 'USD', 2020), ('American Airlines Group Inc', 'Industrails', 'USD', 2021)
Ticker CCK has missing values in Price To Cash Flow Per Share (Daily Time Series Ratio) for the years ('Crown Holdings', 'Materials', 'USD', 2015), ('Crown Holdings', 'Materials', 'USD', 2017), ('Crown Holdings', 'Materials', 'USD', 2018)
Ticker ICAG.L has missing values in Price To Cash Flow Per Share (Daily Time Series Ratio) for the years ('International Consolidated Airlines Group, S.A.', 'Industrails', 'EUR', 2015), ('International Consolidated Airlines Group, S.A.', 'Industrails', 'EUR', 2020), ('International Consolidated Airlines Group, S.A.', 'Industrails', 'EUR', 2021)
Ticker KIE.L has missing values in Price To Cash Flow Per Share (Daily Time Series Ratio) for the years ('Kier Group', 'Industrails', 'GBP', 2015), ('Kier Group', 'Industrails', 'GBP', 2019), ('Kier Group', 'Industrails', 'GBP', 2020)
Ticker 1911.T has missing values in Return On Assets - Actual for the years ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2015), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2016), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2017), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2018), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2019), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2020)
Ticker AJINY.PK has missing values in Return On Assets - Actual for the years ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2015), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2016), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2017), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2018), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2019), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2020), ('Ajinomoto Co.Inc.', 'Consumer Staples', 'JPY', 2021)
Ticker ARDS.AS has missing values in Return On Assets - Actual for the years ('Arcadis', 'Industrails', 'EUR', 2015), ('Arcadis', 'Industrails', 'EUR', 2016), ('Arcadis', 'Industrails', 'EUR', 2021)
Ticker ATEA.OL has missing values in Return On Assets - Actual for the years ('Atea ASA', 'Information Technology', 'NOK', 2015), ('Atea ASA', 'Information Technology', 'NOK', 2016), ('Atea ASA', 'Information Technology', 'NOK', 2017), ('Atea ASA', 'Information Technology', 'NOK', 2018), ('Atea ASA', 'Information Technology', 'NOK', 2019), ('Atea ASA', 'Information Technology', 'NOK', 2020), ('Atea ASA', 'Information Technology', 'NOK', 2021)
Ticker BTI has missing values in Return On Assets - Actual for the years ('British American Tobacco', 'Consumer Staples', 'GBP', 2015), ('British American Tobacco', 'Consumer Staples', 'GBP', 2016), ('British American

Tobacco', 'Consumer Staples', 'GBP', 2021)
Ticker CWK.L has missing values in Return On Assets - Actual for the years ('Cranswick', 'Consumer Staples', 'GBP', 2015), ('Cranswick', 'Consumer Staples', 'GBP', 2016), ('Cranswick', 'Consumer Staples', 'GBP', 2017), ('Cranswick', 'Consumer Staples', 'GBP', 2018), ('Cranswick', 'Consumer Staples', 'GBP', 2020)
Ticker DELL.K has missing values in Return On Assets - Actual for the years ('Dell Technologies', 'Information Technology', 'USD', 2015), ('Dell Technologies', 'Information Technology', 'USD', 2016), ('Dell Technologies', 'Information Technology', 'USD', 2017), ('Dell Technologies', 'Information Technology', 'USD', 2018)
Ticker ESSITYb.ST has missing values in Return On Assets - Actual for the years ('Essity', 'Consumer Staples', 'SEK', 2015), ('Essity', 'Consumer Staples', 'SEK', 2016), ('Essity', 'Consumer Staples', 'SEK', 2017)
Ticker GES has missing values in Return On Assets - Actual for the years ('Guess ?, Inc.', 'Consumer Discretionary', 'USD', 2015), ('Guess ?, Inc.', 'Consumer Discretionary', 'USD', 2020), ('Guess ?, Inc.', 'Consumer Discretionary', 'USD', 2021)
Ticker KIE.L has missing values in Return On Assets - Actual for the years ('Kier Group', 'Industrails', 'GBP', 2015), ('Kier Group', 'Industrails', 'GBP', 2020), ('Kier Group', 'Industrails', 'GBP', 2021)
Ticker KLBN4.SA has missing values in Return On Assets - Actual for the years ('Klabin S/A', 'Materials', 'BRL', 2015), ('Klabin S/A', 'Materials', 'BRL', 2016), ('Klabin S/A', 'Materials', 'BRL', 2018), ('Klabin S/A', 'Materials', 'BRL', 2019), ('Klabin S/A', 'Materials', 'BRL', 2020)
Ticker KNBWY.PK has missing values in Return On Assets - Actual for the years ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2015), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2016), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2017), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2018), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2019), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2020), ('Kirin Holdings Co Ltd', 'Consumer Staples', 'JPY', 2021)
Ticker KSP.I has missing values in Return On Assets - Actual for the years ('Kingspan Group PLC', 'Industrails', 'EUR', 2015), ('Kingspan Group PLC', 'Industrails', 'EUR', 2016), ('Kingspan Group PLC', 'Industrails', 'EUR', 2017)
Ticker LRLCY.PK has missing values in Return On Assets - Actual for the years ("L'Oréal", 'Consumer Staples', 'EUR', 2015), ("L'Oréal", 'Consumer Staples', 'EUR', 2016), ("L'Oréal", 'Consumer Staples', 'EUR', 2017), ("L'Oréal", 'Consumer Staples', 'EUR', 2018), ("L'Oréal", 'Consumer Staples', 'EUR', 2019), ("L'Oréal", 'Consumer Staples', 'EUR', 2020), ("L'Oréal", 'Consumer Staples', 'EUR', 2021)
Ticker PPRUY.PK has missing values in Return On Assets - Actual for the years ('Kering', 'Consumer Discretionary', 'EUR', 2015), ('Kering', 'Consumer Discretionary', 'EUR', 2016), ('Kering', 'Consumer Discretionary', 'EUR', 2017), ('Kering', 'Consumer Discretionary', 'EUR', 2018), ('Kering', 'Consumer Discretionary', 'EUR', 2019), ('Kering', 'Consumer Discretionary', 'EUR', 2020), ('Kering', 'Consumer Discretionary', 'EUR', 2021)
Ticker SBGSY.PK has missing values in Return On Assets - Actual for the years

('Schneider Electric', 'Industrails', 'EUR', 2015), ('Schneider Electric', 'Industrails', 'EUR', 2016), ('Schneider Electric', 'Industrails', 'EUR', 2017), ('Schneider Electric', 'Industrails', 'EUR', 2018), ('Schneider Electric', 'Industrails', 'EUR', 2019), ('Schneider Electric', 'Industrails', 'EUR', 2020), ('Schneider Electric', 'Industrails', 'EUR', 2021)
Ticker SKAb.ST has missing values in Return On Assets - Actual for the years ('Skanska AB', 'Industrails', 'SEK', 2015), ('Skanska AB', 'Industrails', 'SEK', 2016), ('Skanska AB', 'Industrails', 'SEK', 2021)
Ticker SONY.K has missing values in Return On Assets - Actual for the years ('Sony Group Corporation', 'Technology', 'JPY', 2015), ('Sony Group Corporation', 'Technology', 'JPY', 2016), ('Sony Group Corporation', 'Technology', 'JPY', 2017), ('Sony Group Corporation', 'Technology', 'JPY', 2018)
Ticker SOPR.PA has missing values in Return On Assets - Actual for the years ('Sopra Steria Group', 'Information Technology', 'EUR', 2015), ('Sopra Steria Group', 'Information Technology', 'EUR', 2016), ('Sopra Steria Group', 'Information Technology', 'EUR', 2017), ('Sopra Steria Group', 'Information Technology', 'EUR', 2018), ('Sopra Steria Group', 'Information Technology', 'EUR', 2019), ('Sopra Steria Group', 'Information Technology', 'EUR', 2020), ('Sopra Steria Group', 'Information Technology', 'EUR', 2021)
Ticker STN.TO has missing values in Return On Assets - Actual for the years ('Stantec Inc.', 'Industrails', 'CAD', 2015), ('Stantec Inc.', 'Industrails', 'CAD', 2016), ('Stantec Inc.', 'Industrails', 'CAD', 2017), ('Stantec Inc.', 'Industrails', 'CAD', 2019), ('Stantec Inc.', 'Industrails', 'CAD', 2020), ('Stantec Inc.', 'Industrails', 'CAD', 2021)
Ticker VEI.OL has missing values in Return On Assets - Actual for the years ('Veidekke ASA', 'Industrails', 'NOK', 2015), ('Veidekke ASA', 'Industrails', 'NOK', 2017), ('Veidekke ASA', 'Industrails', 'NOK', 2018), ('Veidekke ASA', 'Industrails', 'NOK', 2021)
Ticker VFC has missing values in Return On Assets - Actual for the years ('VF Corporation', 'Consumer Discretionary', 'USD', 2015), ('VF Corporation', 'Consumer Discretionary', 'USD', 2016), ('VF Corporation', 'Consumer Discretionary', 'USD', 2017)
Ticker VMW has missing values in Return On Assets - Actual for the years ('VMware, Inc', 'Information Technology', 'USD', 2015), ('VMware, Inc', 'Information Technology', 'USD', 2016), ('VMware, Inc', 'Information Technology', 'USD', 2017)
Ticker WSP.TO has missing values in Return On Assets - Actual for the years ('WSP', 'Industrails', 'CAD', 2015), ('WSP', 'Industrails', 'CAD', 2016), ('WSP', 'Industrails', 'CAD', 2017), ('WSP', 'Industrails', 'CAD', 2020), ('WSP', 'Industrails', 'CAD', 2021)
Ticker 1911.T has missing values in Total Debt to Total Equity, Percent for the years ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2015), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2016), ('Sumitomo Forestry Co., Ltd.', 'Consumer Discretionary', 'JPY', 2021)
Ticker AAL.O has missing values in Total Debt to Total Equity, Percent for the years ('American Airlines Group Inc', 'Industrails', 'USD', 2015), ('American Airlines Group Inc', 'Industrails', 'USD', 2017), ('American Airlines Group

Inc', 'Industrails', 'USD', 2018), ('American Airlines Group Inc', 'Industrails', 'USD', 2019), ('American Airlines Group Inc', 'Industrails', 'USD', 2020), ('American Airlines Group Inc', 'Industrails', 'USD', 2021)
Ticker ABBV.K has missing values in Total Debt to Total Equity, Percent for the years ('AbbVie Inc', 'Health Care', 'USD', 2015), ('AbbVie Inc', 'Health Care', 'USD', 2018), ('AbbVie Inc', 'Health Care', 'USD', 2019), ('AbbVie Inc', 'Health Care', 'USD', 2021)
Ticker ADSK.O has missing values in Total Debt to Total Equity, Percent for the years ('Autodesk, Inc.', 'Information Technology', 'USD', 2015), ('Autodesk, Inc.', 'Information Technology', 'USD', 2016), ('Autodesk, Inc.', 'Information Technology', 'USD', 2018), ('Autodesk, Inc.', 'Information Technology', 'USD', 2019), ('Autodesk, Inc.', 'Information Technology', 'USD', 2020)
Ticker CL has missing values in Total Debt to Total Equity, Percent for the years ('Colgate Palmolive Company', 'Consumer Staples', 'USD', 2015), ('Colgate Palmolive Company', 'Consumer Staples', 'USD', 2016), ('Colgate Palmolive Company', 'Consumer Staples', 'USD', 2017), ('Colgate Palmolive Company', 'Consumer Staples', 'USD', 2018), ('Colgate Palmolive Company', 'Consumer Staples', 'USD', 2021)
Ticker DELL.K has missing values in Total Debt to Total Equity, Percent for the years ('Dell Technologies', 'Information Technology', 'USD', 2015), ('Dell Technologies', 'Information Technology', 'USD', 2016), ('Dell Technologies', 'Information Technology', 'USD', 2019), ('Dell Technologies', 'Information Technology', 'USD', 2020)
Ticker EONGn.DE has missing values in Total Debt to Total Equity, Percent for the years ('E.ON SE', 'Utilities', 'EUR', 2015), ('E.ON SE', 'Utilities', 'EUR', 2016), ('E.ON SE', 'Utilities', 'EUR', 2021)
Ticker GSK has missing values in Total Debt to Total Equity, Percent for the years ('GSK', 'Health Care', 'GBP', 2015), ('GSK', 'Health Care', 'GBP', 2017), ('GSK', 'Health Care', 'GBP', 2021)
Ticker HPQ has missing values in Total Debt to Total Equity, Percent for the years ('HP Inc', 'Information Technology', 'USD', 2015), ('HP Inc', 'Information Technology', 'USD', 2016), ('HP Inc', 'Information Technology', 'USD', 2017), ('HP Inc', 'Information Technology', 'USD', 2018), ('HP Inc', 'Information Technology', 'USD', 2019), ('HP Inc', 'Information Technology', 'USD', 2020), ('HP Inc', 'Information Technology', 'USD', 2021)
Ticker JNJ has missing values in Total Debt to Total Equity, Percent for the years ('Johnson & Johnson', 'Health Care', 'USD', 2015), ('Johnson & Johnson', 'Health Care', 'USD', 2016), ('Johnson & Johnson', 'Health Care', 'USD', 2020)
Ticker PM has missing values in Total Debt to Total Equity, Percent for the years ('Philip Morris International', 'Consumer Staples', 'USD', 2015), ('Philip Morris International', 'Consumer Staples', 'USD', 2016), ('Philip Morris International', 'Consumer Staples', 'USD', 2017), ('Philip Morris International', 'Consumer Staples', 'USD', 2018), ('Philip Morris International', 'Consumer Staples', 'USD', 2019), ('Philip Morris International', 'Consumer Staples', 'USD', 2020), ('Philip Morris International', 'Consumer Staples', 'USD', 2021)
Ticker SBUX.O has missing values in Total Debt to Total Equity, Percent for the years ('Starbucks Corporation', 'Consumer Discretionary', 'USD', 2015),

```
('Starbucks Corporation', 'Consumer Discretionary', 'USD', 2016), ('Starbucks
Corporation', 'Consumer Discretionary', 'USD', 2019), ('Starbucks Corporation',
'Consumer Discretionary', 'USD', 2020), ('Starbucks Corporation', 'Consumer
Discretionary', 'USD', 2021)
Ticker YUM has missing values in Total Debt to Total Equity, Percent for the
years ('Yum! Brands, Inc.', 'Consumer Discretionary', 'USD', 2015), ('Yum!
Brands, Inc.', 'Consumer Discretionary', 'USD', 2016), ('Yum! Brands, Inc.',
'Consumer Discretionary', 'USD', 2017), ('Yum! Brands, Inc.', 'Consumer
Discretionary', 'USD', 2018), ('Yum! Brands, Inc.', 'Consumer Discretionary',
'USD', 2019), ('Yum! Brands, Inc.', 'Consumer Discretionary', 'USD', 2020),
('Yum! Brands, Inc.', 'Consumer Discretionary', 'USD', 2021)
Ticker CSCO.O has missing values in Weighted Average Cost of Capital, (%) for
the years ('Cisco Systems, Inc.', 'Information Technology', 'USD', 2015),
('Cisco Systems, Inc.', 'Information Technology', 'USD', 2016), ('Cisco Systems,
Inc.', 'Information Technology', 'USD', 2017), ('Cisco Systems, Inc.',
'Information Technology', 'USD', 2018), ('Cisco Systems, Inc.', 'Information
Technology', 'USD', 2019), ('Cisco Systems, Inc.', 'Information Technology',
'USD', 2020)
Ticker EONGn.DE has missing values in Weighted Average Cost of Capital, (%) for
the years ('E.ON SE', 'Utilities', 'EUR', 2015), ('E.ON SE', 'Utilities', 'EUR',
2016), ('E.ON SE', 'Utilities', 'EUR', 2017)
Ticker HMb.ST has missing values in Weighted Average Cost of Capital, (%) for
the years ('H&M Hennes & Mauritz AB', 'Consumer Discretionary', 'SEK', 2015),
('H&M Hennes & Mauritz AB', 'Consumer Discretionary', 'SEK', 2017), ('H&M Hennes
& Mauritz AB', 'Consumer Discretionary', 'SEK', 2018)
Ticker NG.L has missing values in Weighted Average Cost of Capital, (%) for the
years ('National Grid PLC', 'Utilities', 'GBP', 2015), ('National Grid PLC',
'Utilities', 'GBP', 2019), ('National Grid PLC', 'Utilities', 'GBP', 2020)
```

[26]:
```
count_no_missing_values(df_convert, metrics)
```

```
61 tickers have no missing values in Asset Turnover for the years 2017-2021.
64 tickers have no missing values in Current Ratio for the years 2017-2021.
173 tickers have no missing values in Enterprise Value To EBIT (Daily Time
Series Ratio) for the years 2017-2021.
57 tickers have no missing values in Interest Coverage Ratio for the years
2017-2021.
173 tickers have no missing values in Price To Cash Flow Per Share (Daily Time
Series Ratio) for the years 2017-2021.
163 tickers have no missing values in Return On Assets - Actual for the years
2017-2021.
60 tickers have no missing values in Total Debt to Total Equity, Percent for the
years 2017-2021.
185 tickers have no missing values in Weighted Average Cost of Capital, (%) for
the years 2017-2021.
```

[27]:
```
def get_no_missing_values_tickers(Fin_df, metrics):
    no_missing_values_tickers = []

    # Loop through each ticker
    for ticker in Fin_df.index.get_level_values('Ticker').unique():
        # Get the data for the current ticker
        ticker_data = Fin_df.loc[ticker]

        # Filter for the years 2016-2020
        ticker_data = ticker_data.loc[(ticker_data.index.
 ↪get_level_values('Year') >= 2017) & (ticker_data.index.
 ↪get_level_values('Year') <= 2020)]

        # Check if there are any missing values for any of the metrics in these␣
 ↪years
        if all(ticker_data[metric].isnull().sum() == 0 for metric in metrics):
            no_missing_values_tickers.append(ticker)

    return no_missing_values_tickers
```

[28]:
```
# Get the list of tickers with no missing values for these metrics in the years␣
 ↪2016-2020
tickers = get_no_missing_values_tickers(df_convert, metrics)

df_filtered = df_convert.loc[tickers]
df_filtered = df_filtered[(df_filtered.index.get_level_values('Year') >= 2017)␣
 ↪& (df_filtered.index.get_level_values('Year') <= 2020)]
```

[29]:
```
df_filtered.reset_index(inplace=True, drop=False)
df_filtered.drop(['Name','Currency'],axis=1,inplace=True)

##########################################################
# Converting Year to Year+1 to merge with emissions data #
##########################################################
df_filtered['Year'] = df_filtered['Year'] + 1
df_filtered.set_index(['Ticker','Sector','Year'],inplace=True)
df_filtered
```

[29]:
```
Metric                                Asset Turnover Current Ratio  \
Ticker Sector                 Year
1925.T Real Estate            2018       1.031133       1.37498
                              2019       1.000137       1.44366
                              2020       0.990192       1.37036
                              2021        0.97757       1.62178
1928.T Consumer Discretionary 2018       0.961841       1.89795
…                                            …            …
WMT    Consumer Staples       2021       2.299147       0.79452
```

```
XYL    Industrails            2018       0.706015       1.88273
                              2019       0.739526       1.50756
                              2020       0.703054       1.63225
                              2021       0.592467       1.80112

Metric                              Enterprise Value To EBIT (Daily Time Series
Ratio)  \
Ticker Sector                 Year
1925.T Real Estate            2018
9.892028
                              2019
8.199919
                              2020
7.747042
                              2021
9.482526
1928.T Consumer Discretionary 2018
10.223925
…
…
WMT    Consumer Staples       2021
17.39986
XYL    Industrails            2018
27.439846
                              2019
21.509368
                              2020
18.654773
                              2021
33.352629

Metric                              Interest Coverage Ratio  \
Ticker Sector                 Year
1925.T Real Estate            2018               60.293992
                              2019                62.61562
                              2020               49.599547
                              2021               42.430862
1928.T Consumer Discretionary 2018               79.176268
…                                                     …
WMT    Consumer Staples       2021                8.067718
XYL    Industrails            2018                7.036585
                              2019                8.243902
                              2020               12.970149
                              2021                 7.61039

Metric                              Price To Cash Flow Per Share (Daily Time
Series Ratio)  \
```

```
Ticker Sector                 Year
1925.T Real Estate            2018
10.020518
                              2019
6.094981
                              2020
6.351389
                              2021
13.640672
1928.T Consumer Discretionary 2018
12.138439
…
…
WMT    Consumer Staples       2021
12.139587
XYL    Industrails            2018
20.346711
                              2019
17.252741
                              2020
21.861963
                              2021
21.788378

Metric                             Return On Assets - Actual  \
Ticker Sector                 Year
1925.T Real Estate            2018                       8.8
                              2019                       9.1
                              2020                       8.6
                              2021                       8.2
1928.T Consumer Discretionary 2018                       9.1
…                                                        …
WMT    Consumer Staples       2021                       6.2
XYL    Industrails            2018                      6.54
                              2019                      6.49
                              2020                      3.34
                              2021                      7.33

Metric                             Total Debt to Total Equity, Percent  \
Ticker Sector                 Year
1925.T Real Estate            2018                           51.015434
                              2019                           55.444592
                              2020                           52.272561
                              2021                           65.935767
1928.T Consumer Discretionary 2018                           44.473771
…                                                                  …
WMT    Consumer Staples       2021                           72.947274
```

```
XYL    Industrails            2018                            87.894527
                              2019                            83.381503
                              2020                            78.322624
                              2021                           103.908356

Metric                              Weighted Average Cost of Capital, (%)
Ticker Sector                 Year
1925.T Real Estate            2018                                4.789896
                              2019                                4.564808
                              2020                                2.684953
                              2021                                2.188228
1928.T Consumer Discretionary 2018                                3.908878
…                                                                    …
WMT    Consumer Staples       2021                                3.210361
XYL    Industrails            2018                                7.401344
                              2019                                 9.34875
                              2020                                7.666595
                              2021                                6.079927

[512 rows x 8 columns]
```

[30]:
```
## There are 135 companies available with no missing values for these metrics␣
↪in the years 2016-2020
len(df_filtered.index.get_level_values('Ticker').unique())
```

[30]: 128

[31]:
```
## Consider this new features as they have new names
df_filtered.columns
```

[31]:
```
Index(['Asset Turnover', 'Current Ratio',
       'Enterprise Value To EBIT (Daily Time Series Ratio)',
       'Interest Coverage Ratio',
       'Price To Cash Flow Per Share (Daily Time Series Ratio)',
       'Return On Assets - Actual', 'Total Debt to Total Equity, Percent',
       'Weighted Average Cost of Capital, (%)'],
      dtype='object', name='Metric')
```

1.3 Final Data: Emission Data and Financial Data put together

[32]:
```
final_data = pd.merge(data, df_filtered, how='inner',␣
↪on=['Ticker','Sector','Year'])
final_data.reset_index(['Sector'],inplace=True)
final_data
```

[32]:

Ticker	Year	Sector	Scope1	Scope2 Location	Scope2 Market	\
ABBNY.PK	2018	Industrails	375000.0	650799.0	627903.0	
	2019	Industrails	257500.0	456000.0	410800.0	
	2020	Industrails	225547.0	414000.0	335300.0	
	2021	Industrails	190900.0	370000.0	214000.0	
AENA.MC	2018	Industrails	30707.0	236695.0	218937.0	
…		…	…	…	…	
XYL	2021	Industrails	47707.0	44569.0	18214.0	
ORSTED.CO	2018	Utilities	3483000.0	151000.0	46000.0	
	2019	Utilities	1846000.0	123000.0	4000.0	
	2020	Utilities	1851000.0	111000.0	1790.0	
	2021	Utilities	2142000.0	53000.0	802.0	

Ticker	Year	Sum_Emission_Mkt	Sum_Emission_Loc	Asset Turnover	\
ABBNY.PK	2018	1002903.0	1025799.0	0.60963	
	2019	668300.0	713500.0	0.629404	
	2020	560847.0	639547.0	0.617964	
	2021	404900.0	560900.0	0.599431	
AENA.MC	2018	249644.0	267402.0	0.257004	
…		…	…	…	
XYL	2021	65921.0	92276.0	0.592467	
ORSTED.CO	2018	3529000.0	3634000.0	0.421957	
	2019	1850000.0	1969000.0	0.470389	
	2020	1852790.0	1962000.0	0.383186	
	2021	2142802.0	2195000.0	0.257463	

Ticker	Year	Current Ratio	\
ABBNY.PK	2018	1.34404	
	2019	1.18529	
	2020	1.30554	
	2021	1.24715	
AENA.MC	2018	0.81773	
…		…	
XYL	2021	1.80112	
ORSTED.CO	2018	2.09084	
	2019	1.90445	
	2020	2.03417	
	2021	1.83631	

Ticker	Year	Enterprise Value To EBIT (Daily Time Series Ratio)	\
ABBNY.PK	2018	19.555091	
	2019	19.925733	
	2020	22.149453	

	2021	19.09039
AENA.MC	2018	19.303482
…		…
XYL	2021	33.352629
ORSTED.CO	2018	20.717344
	2019	40.983505
	2020	22.119288
	2021	51.580208

		Interest Coverage Ratio \
Ticker	Year	
ABBNY.PK	2018	9.897638
	2019	8.157377
	2020	26.172043
	2021	5.957589
AENA.MC	2018	17.072434
…		…
XYL	2021	7.61039
ORSTED.CO	2018	11.374454
	2019	7.250831
	2020	6.920519
	2021	4.6374

		Price To Cash Flow Per Share (Daily Time Series Ratio) \
Ticker	Year	
ABBNY.PK	2018	16.3461
	2019	13.280204
	2020	21.598579
	2021	23.154692
AENA.MC	2018	12.940817
…		…
XYL	2021	21.788378
ORSTED.CO	2018	15.257509
	2019	31.059867
	2020	17.665441
	2021	33.70584

		Return On Assets - Actual Total Debt to Total Equity, Percent \	
Ticker	Year		
ABBNY.PK	2018	6.88	49.989878
	2019	5.736	61.768922
	2020	6.46	66.974715
	2021	5.85	39.024546
AENA.MC	2018	7.156	138.647566
…		…	…
XYL	2021	7.33	103.908356
ORSTED.CO	2018	8.29	43.563134

```
          2019                      8.89                          33.398999
          2020                      0.86                          48.858818
          2021                      3.54                           44.20345

                Weighted Average Cost of Capital, (%)
Ticker    Year
ABBNY.PK  2018                              6.604438
          2019                              7.768313
          2020                              7.130816
          2021                              7.200423
AENA.MC   2018                              4.298168
…                                              …
XYL       2021                              6.079927
ORSTED.CO 2018                              4.466487
          2019                              5.118872
          2020                              4.403494
          2021                              4.628572

[512 rows x 14 columns]
```

[33]:
```
final_data.columns = ['Sector','Scope1','Scope2 Location','Scope2␣
↪Market','Sum_Emission_Mkt','Sum_Emission_Loc',
                      'Asset Turnover', 'Current Ratio', 'Enterprise Value to␣
↪EBIT', 'Interest Coverage Ratio',
                      'Price to Cash Flow per Share', 'Return on Assets' ,␣
↪'Total Debt to Total Equity',
                      'Weighted Average Cost of Capital']
```

[34]:
```
for x in final_data.columns[1:]:
    final_data[x]=final_data[x].astype(float)
final_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
MultiIndex: 512 entries, ('ABBNY.PK', 2018) to ('ORSTED.CO', 2021)
Data columns (total 14 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Sector                          512 non-null    object
 1   Scope1                          512 non-null    float64
 2   Scope2 Location                 512 non-null    float64
 3   Scope2 Market                   512 non-null    float64
 4   Sum_Emission_Mkt                512 non-null    float64
 5   Sum_Emission_Loc                512 non-null    float64
 6   Asset Turnover                  512 non-null    float64
 7   Current Ratio                   512 non-null    float64
 8   Enterprise Value to EBIT        512 non-null    float64
 9   Interest Coverage Ratio         512 non-null    float64
```

```
 10  Price to Cash Flow per Share      512 non-null    float64
 11  Return on Assets                  512 non-null    float64
 12  Total Debt to Total Equity        512 non-null    float64
 13  Weighted Average Cost of Capital  512 non-null    float64
dtypes: float64(13), object(1)
memory usage: 62.9+ KB
```

[35]:
```
len(final_data.index.get_level_values('Ticker').unique())
```

[35]:
```
128
```

[36]:
```
final_data.index.get_level_values('Year').unique()
```

[36]:
```
Int64Index([2018, 2019, 2020, 2021], dtype='int64', name='Year')
```

2 Exploratory Data Analysis and Visualisation

2.1 Descriptive Statistics

Calculate summary statistics for each metric

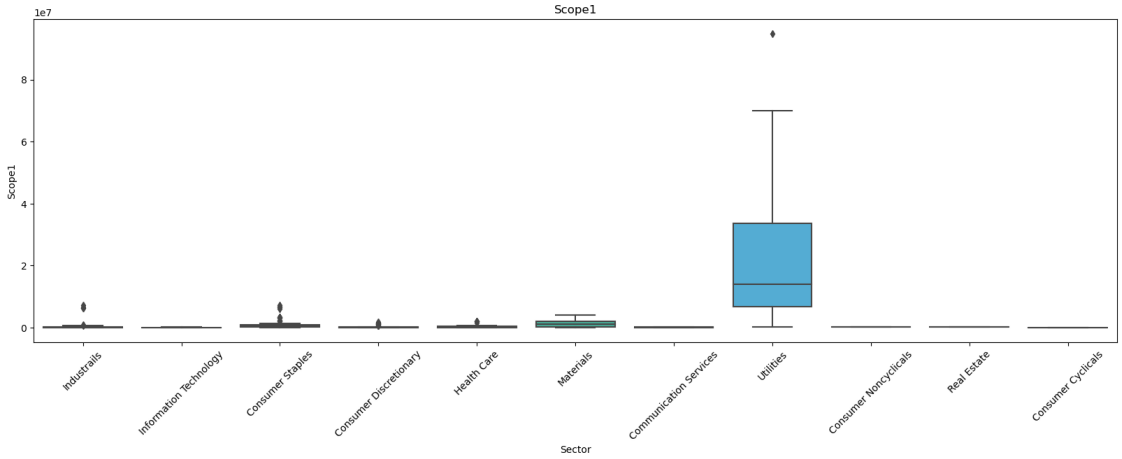
[37]:
```
final_data.describe().round(2)
```

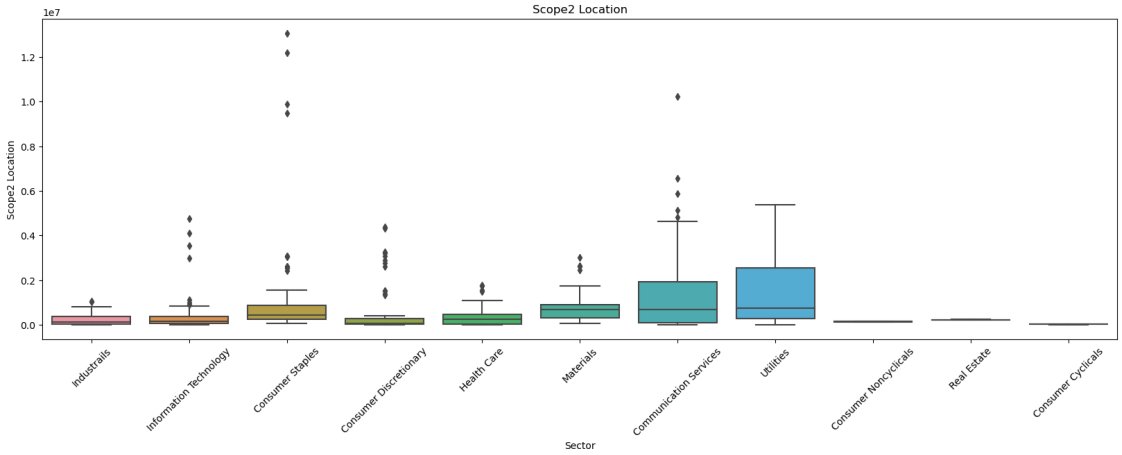
[37]:
```
            Scope1  Scope2 Location  Scope2 Market  Sum_Emission_Mkt  \
count       512.00           512.00         512.00            512.00
mean    2069802.02        747378.48      534023.05        2603825.07
std     8137275.12       1467282.31     1190225.15        8705552.63
min          28.00          2347.00          61.00           2211.00
25%       25535.50         62834.75       28175.00          91849.00
50%      155940.00        267403.50      155882.50         337116.50
75%      608561.25        730836.25      502587.25        1210011.00
max    94794693.00      13057352.00    12022083.00       96415272.00

       Sum_Emission_Loc  Asset Turnover  Current Ratio  \
count            512.00          512.00         512.00
mean         2817180.50            0.80           1.42
std          8701118.89            0.51           0.71
min             4055.00            0.08           0.36
25%           128668.75            0.46           0.94
50%           447343.50            0.67           1.26
75%          1592783.00            0.98           1.70
max         95884305.00            3.30           5.58

       Enterprise Value to EBIT  Interest Coverage Ratio  \
count                    512.00                   512.00
mean                      28.49                    51.51
std                       72.50                   313.06
```

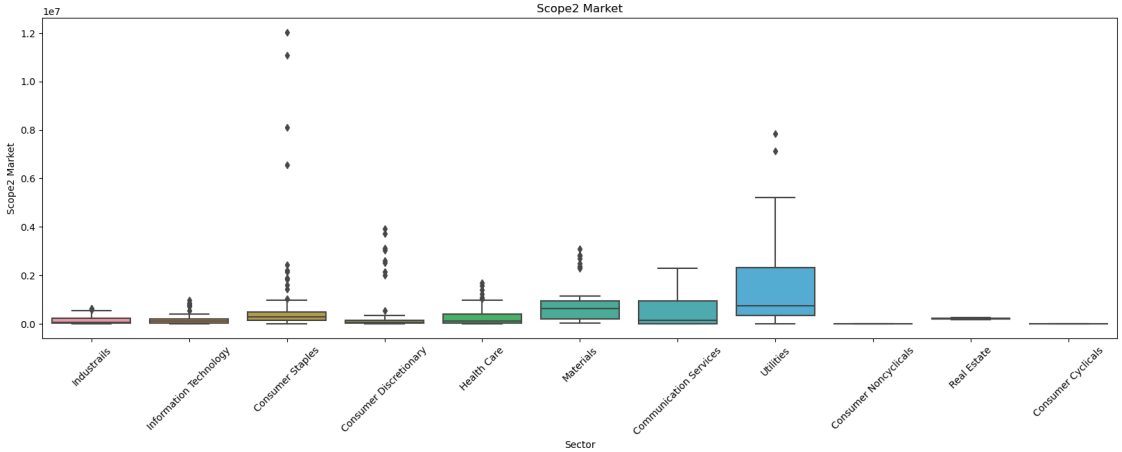
```
min                        4.48                    -5.29
25%                       12.78                     5.62
50%                       16.97                     9.87
75%                       22.64                    24.07
max                      937.94                  5073.53

       Price to Cash Flow per Share  Return on Assets  \
count                        512.00            512.00
mean                          14.34              7.36
std                            9.68              5.20
min                            1.36             -9.44
25%                            7.33              4.27
50%                           12.45              6.50
75%                           19.07              9.02
max                           75.02             40.06

       Total Debt to Total Equity  Weighted Average Cost of Capital
count                      512.00                            512.00
mean                       100.87                              5.87
std                        123.27                              2.05
min                          0.14                              1.05
25%                         44.32                              4.42
50%                         71.05                              5.76
75%                        126.85                              7.38
max                       2032.34                             13.28
```

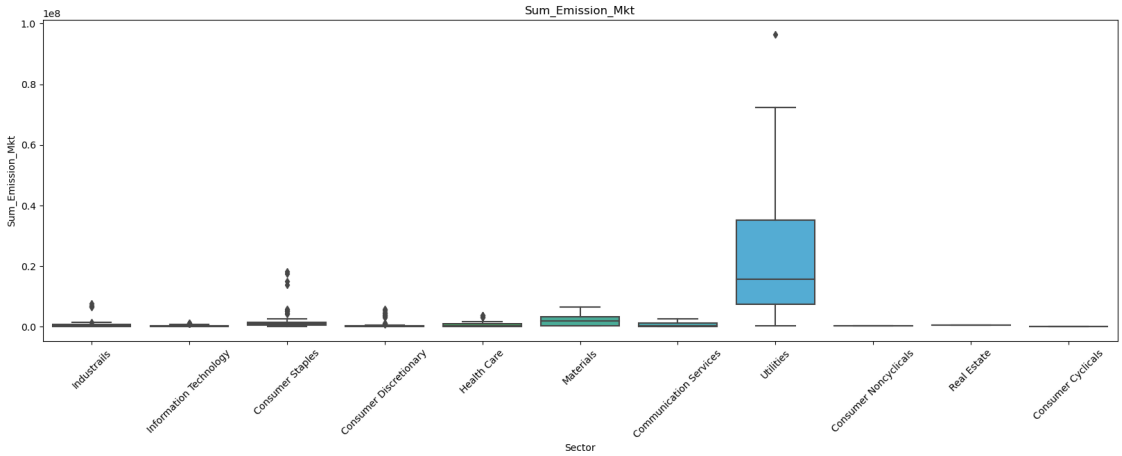
Box Plots by Sector Show distribution of metrics within each sector and highlight outliers

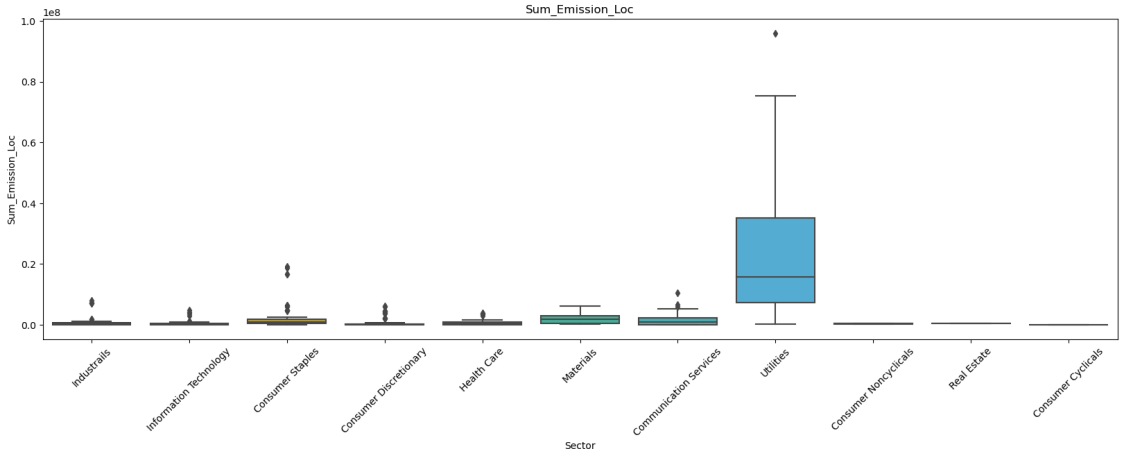
[38]:
```
for metric in final_data.columns[1:]:
    plt.figure(figsize=(20,6))
    sns.boxplot(x='Sector', y=metric, data=final_data.reset_index())
    plt.title(metric)
    plt.xticks(rotation=45)
    plt.show()
```

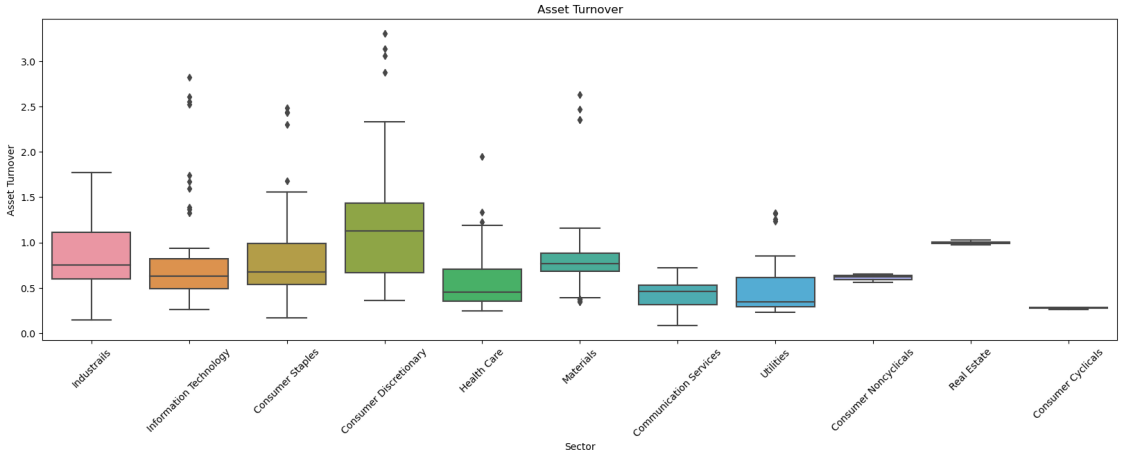

Scope1
1e7
Scope1
Sector
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals

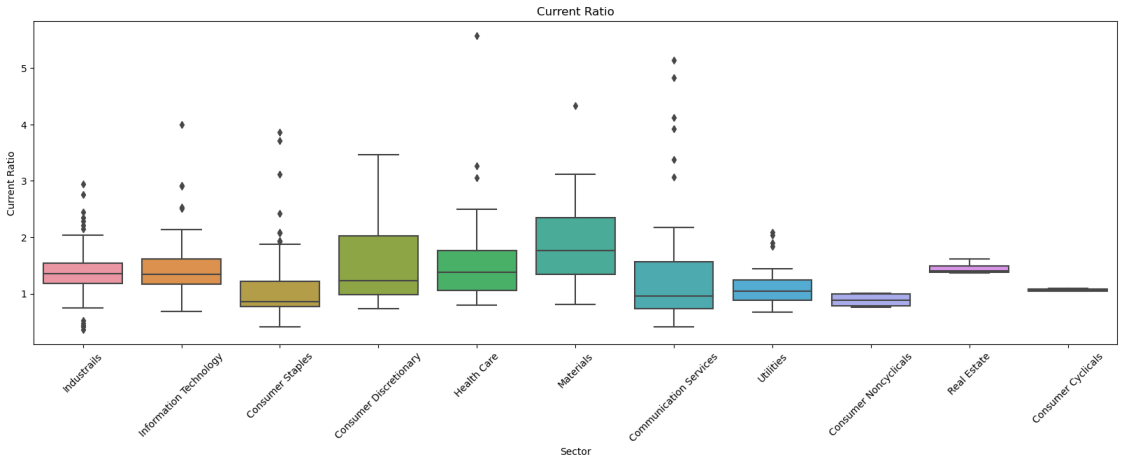

Scope2 Location
1e7
Scope2 Location
Sector
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals

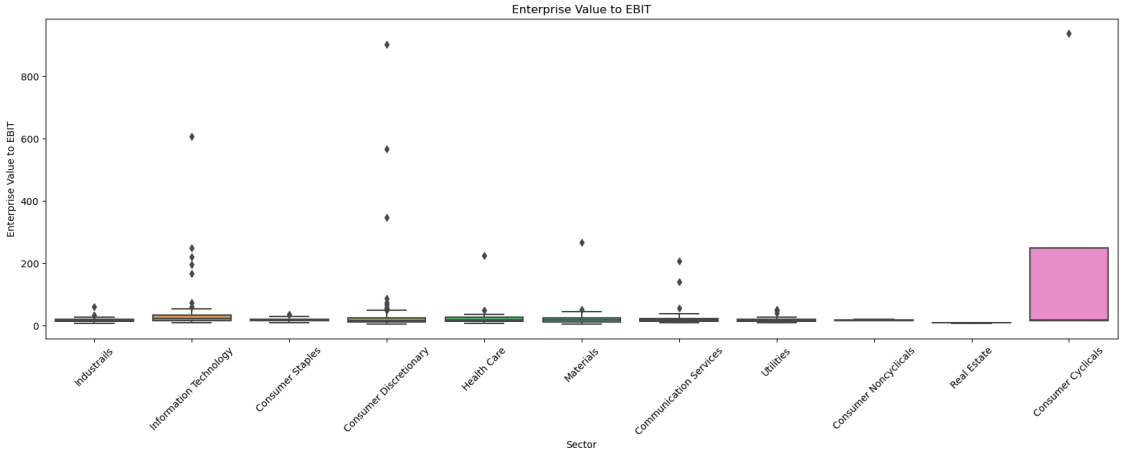

Scope2 Market
1e7
Scope2 Market
Sector
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals

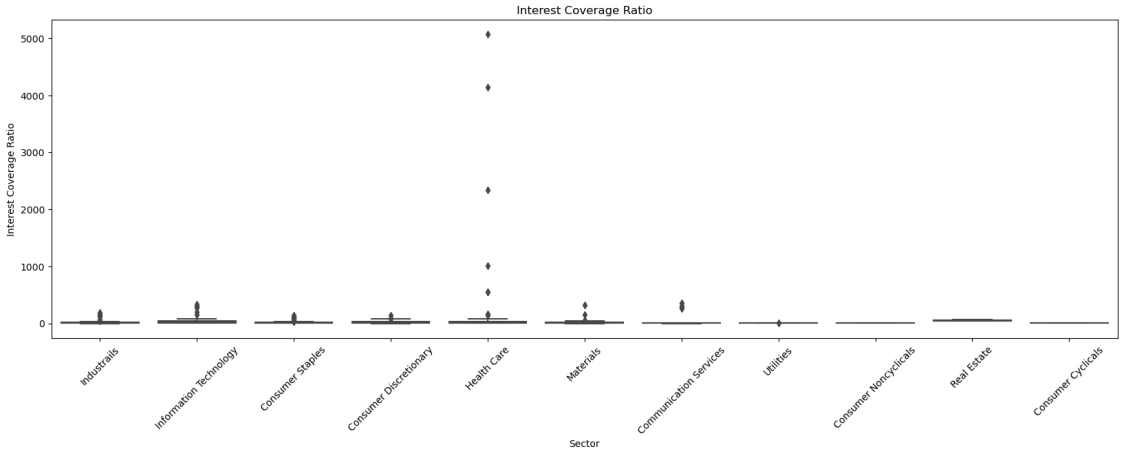

Sum_Emission_Mkt
1e8
1.0
0.8
0.6
0.4
0.2
0.0
Sum_Emission_Mkt
Industrails
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

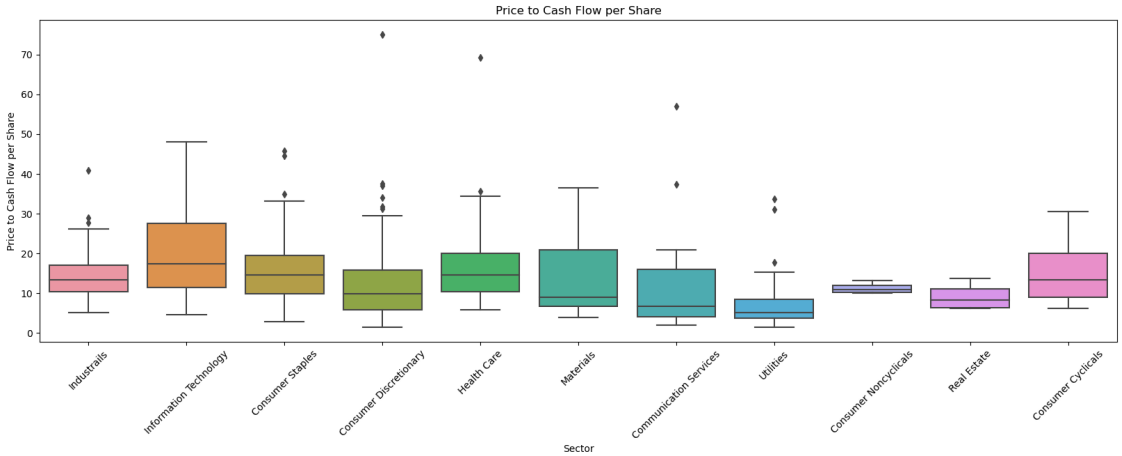

Sum_Emission_Loc
1e8
1.0
0.8
0.6
0.4
0.2
0.0
Sum_Emission_Loc
Industrails
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

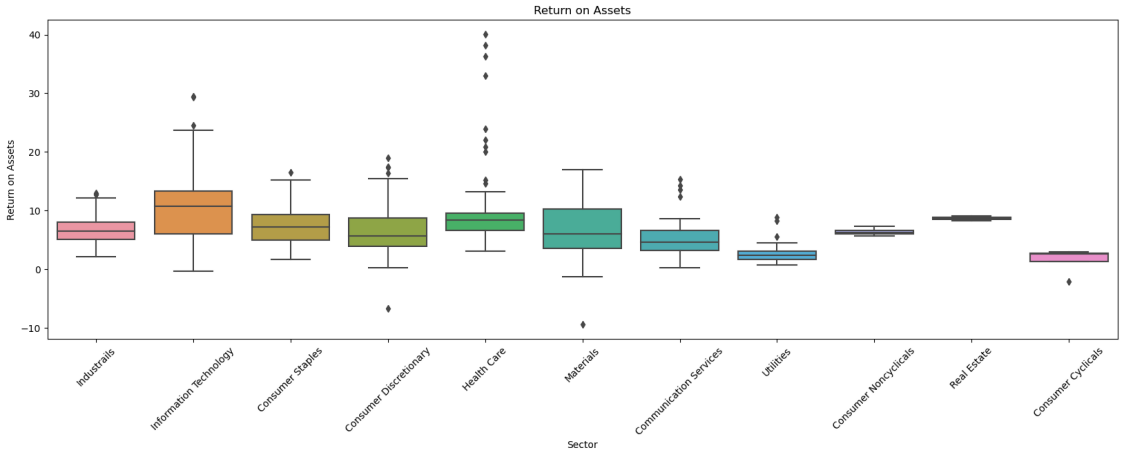

Asset Turnover
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Asset Turnover
Industrails
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

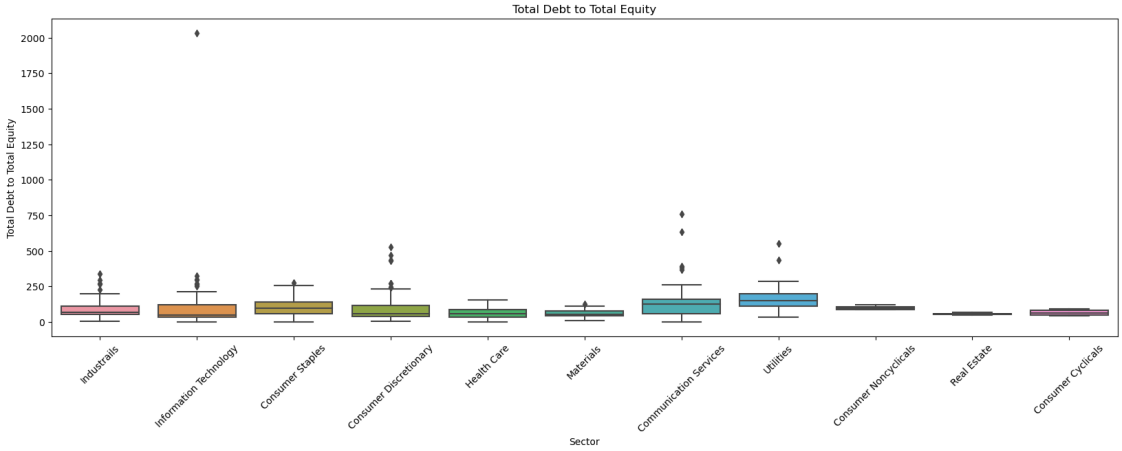

Current Ratio
Current Ratio
1
2
3
4
5
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

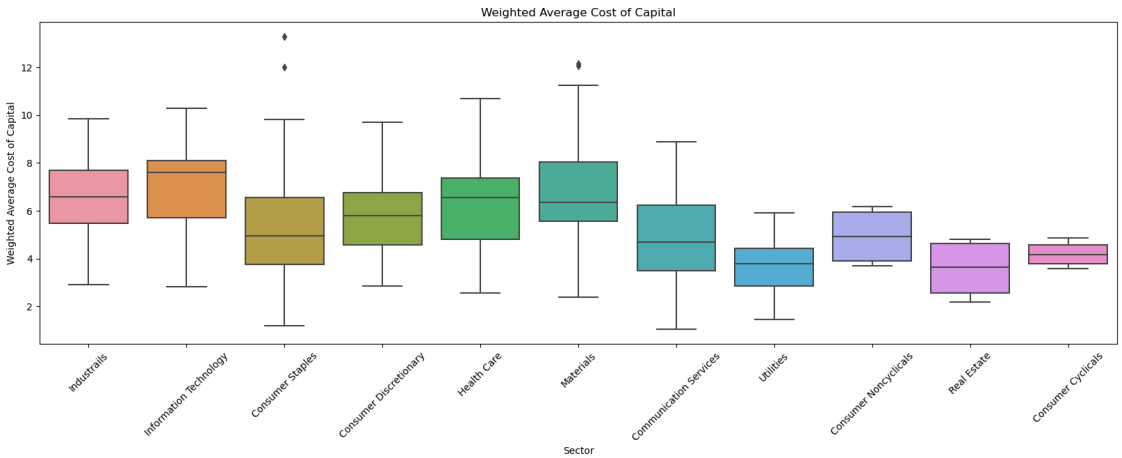

Enterprise Value to EBIT
Enterprise Value to EBIT
0
200
400
600
800
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector


Interest Coverage Ratio
Interest Coverage Ratio
0
1000
2000
3000
4000
5000
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector


Price to Cash Flow per Share
Price to Cash Flow per Share
0
10
20
30
40
50
60
70
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

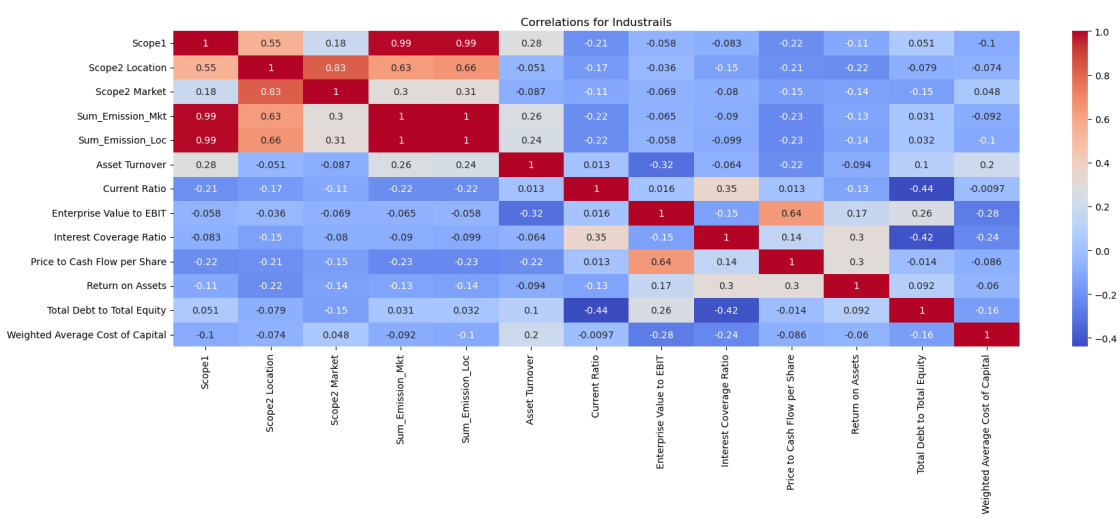

Return on Assets
Return on Assets
−10
0
10
20
30
40
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

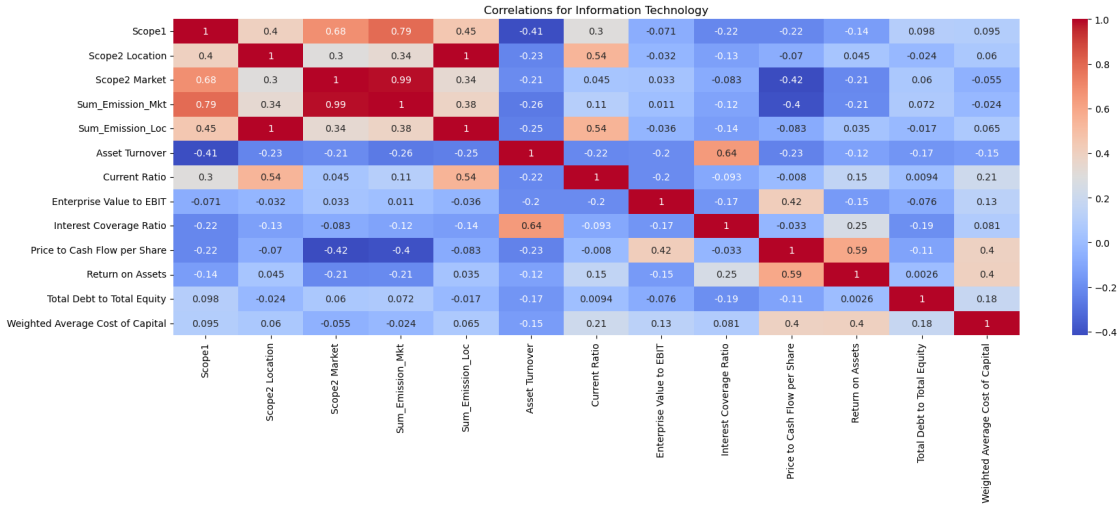

Total Debt to Total Equity
Total Debt to Total Equity
0
250
500
750
1000
1250
1500
1750
2000
Industrials
Information Technology
Consumer Staples
Consumer Discretionary
Health Care
Materials
Communication Services
Utilities
Consumer Noncyclicals
Real Estate
Consumer Cyclicals
Sector

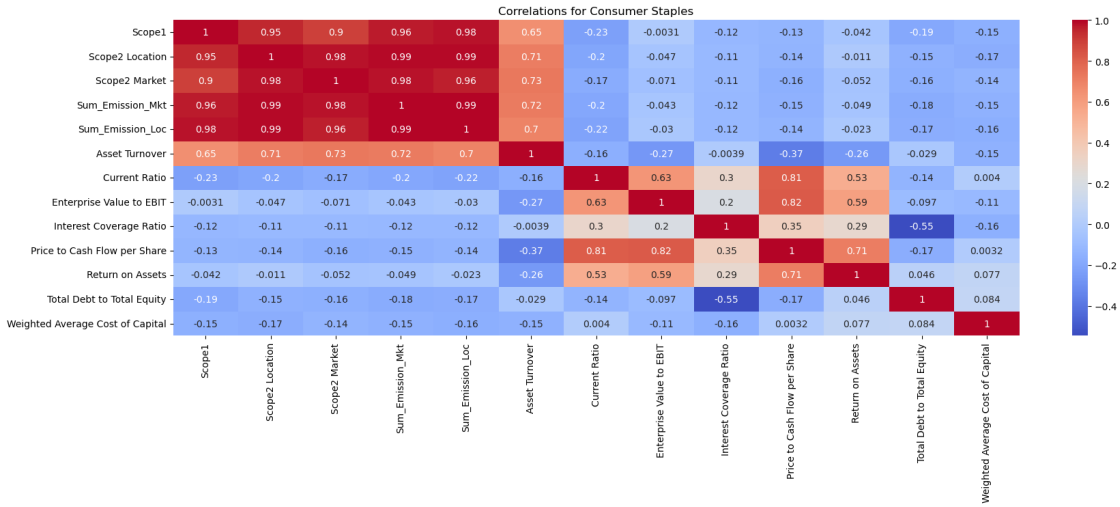


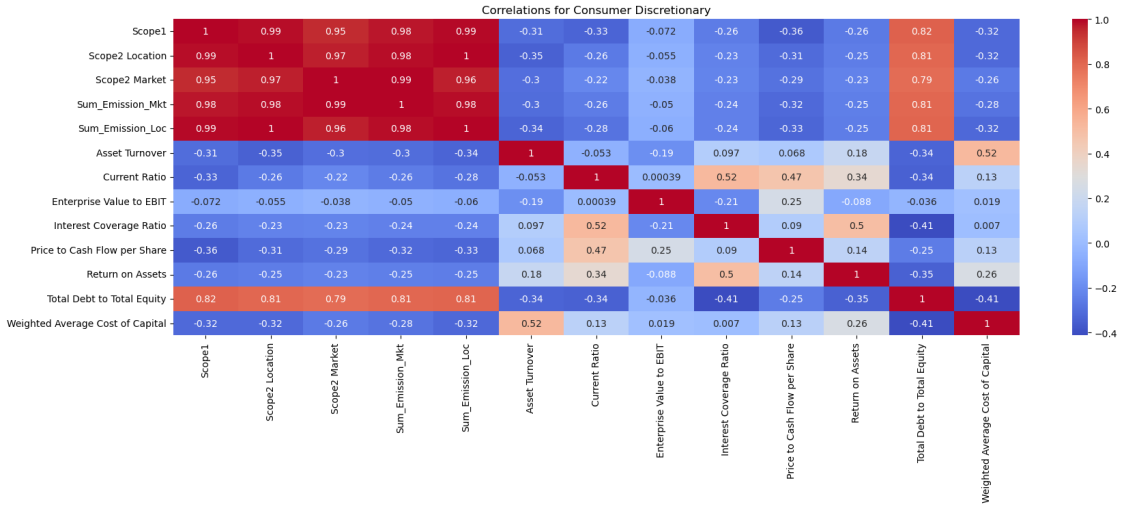
Heatmaps of Correlations Plot a heatmap of the correlations for each sector

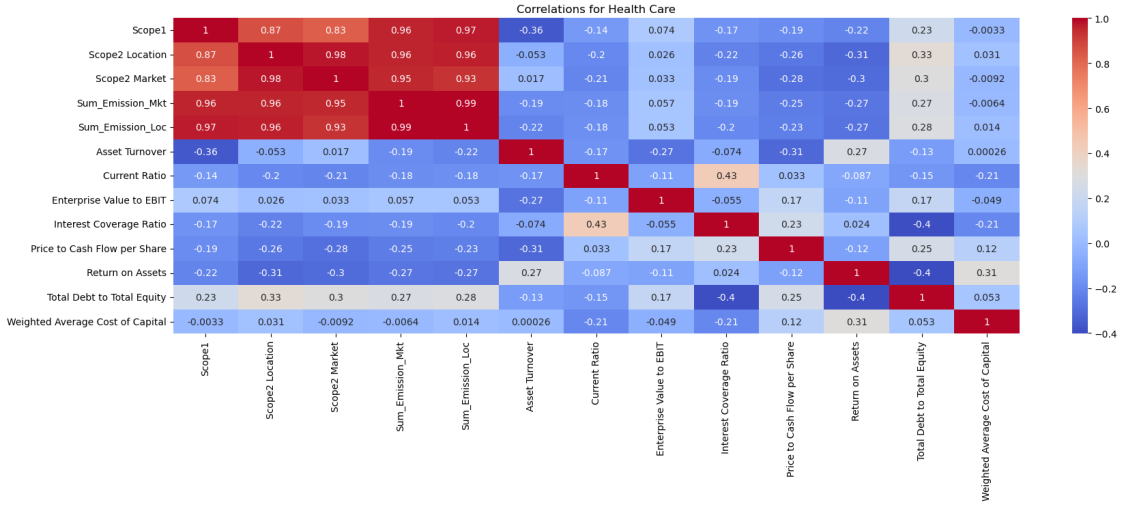
[39]:
```
sectors = final_data.Sector.unique()
correlations_by_sector = {sector: final_data.loc[final_data.Sector ==␣
↪sector][final_data.columns[1:]].corr() for sector in sectors}
for sector, correlations in correlations_by_sector.items():
    plt.figure(figsize=(20,6))
    sns.heatmap(correlations, annot=True, cmap='coolwarm')
    plt.title(f'Correlations for {sector}')
    plt.show()
```

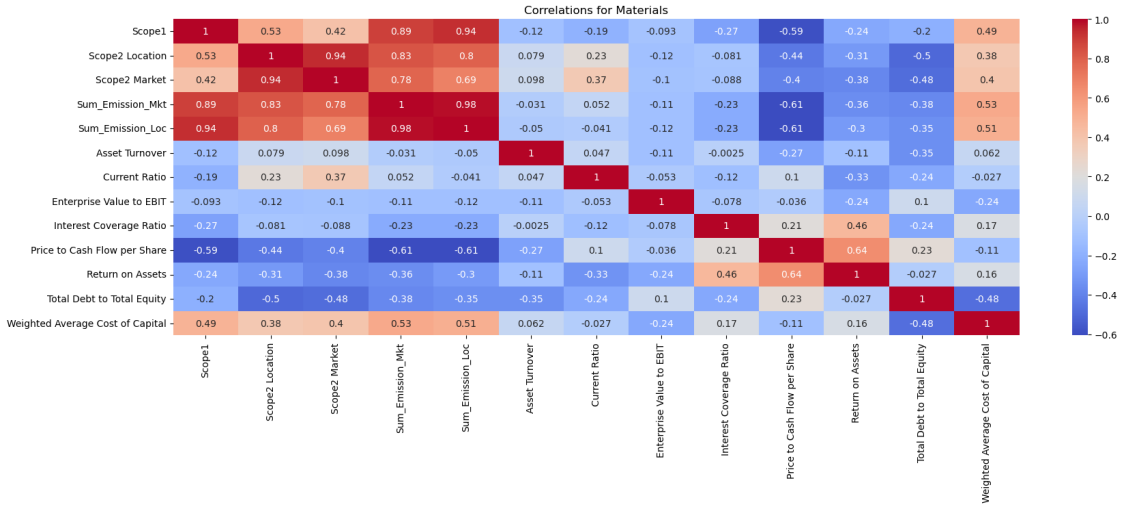


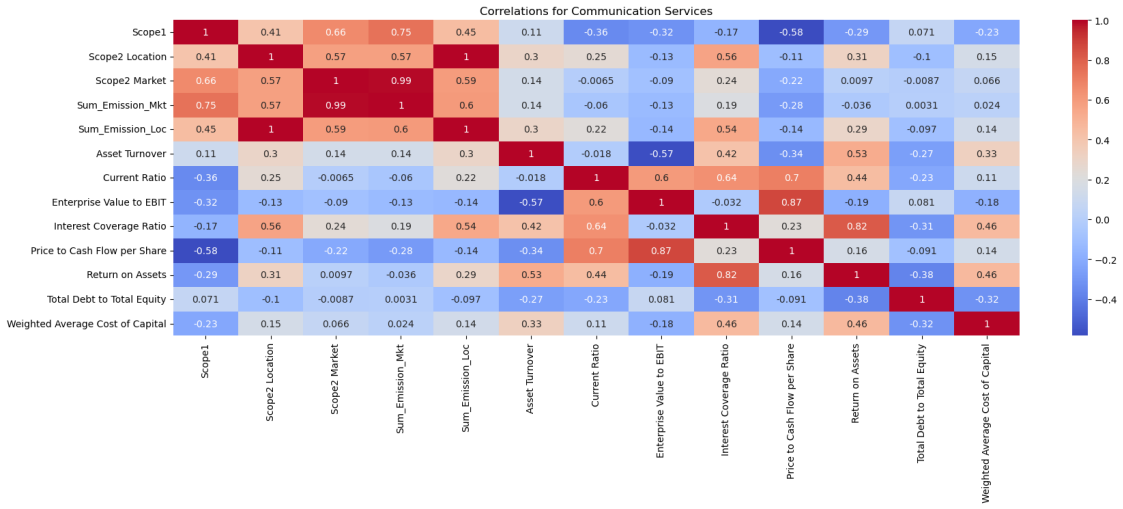

Correlations for Information Technology
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

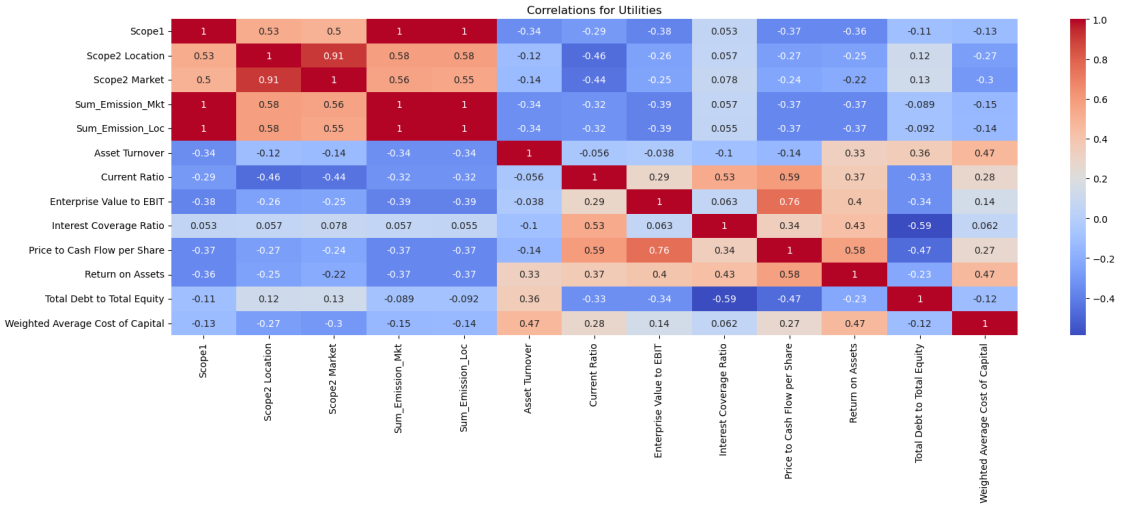

Correlations for Consumer Staples
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

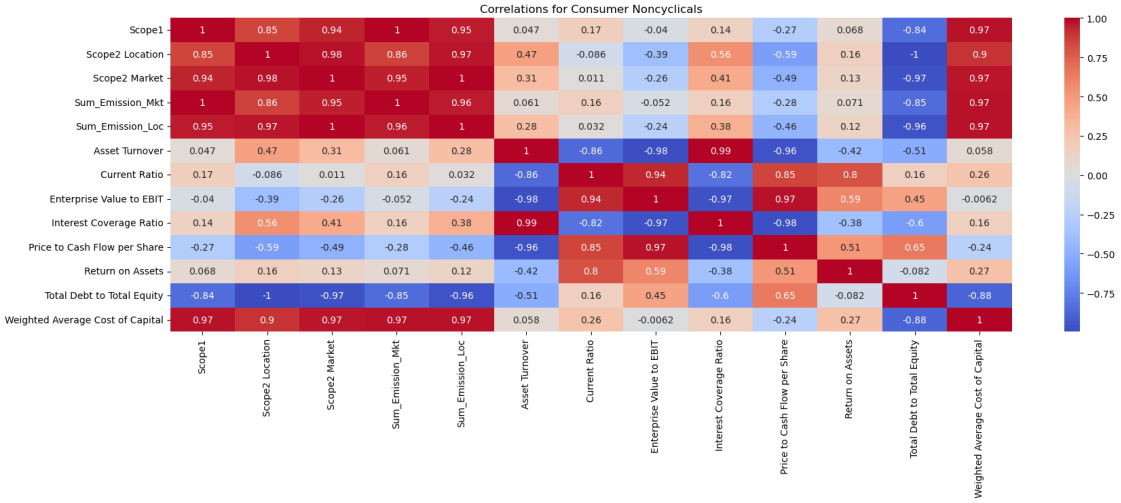

Correlations for Consumer Discretionary
Scope1 1 0.99 0.95 0.98 0.99 -0.31 -0.33 -0.072 -0.26 -0.36 -0.26 0.82 -0.32
Scope2 Location 0.99 1 0.97 0.98 1 -0.35 -0.26 -0.055 -0.23 -0.31 -0.25 0.81 -0.32
Scope2 Market 0.95 0.97 1 0.99 0.96 -0.3 -0.22 -0.038 -0.23 -0.29 -0.23 0.79 -0.26
Sum_Emission_Mkt 0.98 0.98 0.99 1 0.98 -0.3 -0.26 -0.05 -0.24 -0.32 -0.25 0.81 -0.28
Sum_Emission_Loc 0.99 1 0.96 0.98 1 -0.34 -0.28 -0.06 -0.24 -0.33 -0.25 0.81 -0.32
Asset Turnover -0.31 -0.35 -0.3 -0.3 -0.34 1 -0.053 -0.19 0.097 0.068 0.18 -0.34 0.52
Current Ratio -0.33 -0.26 -0.22 -0.26 -0.28 -0.053 1 0.00039 0.52 0.47 0.34 -0.34 0.13
Enterprise Value to EBIT -0.072 -0.055 -0.038 -0.05 -0.06 -0.19 0.00039 1 -0.21 0.25 -0.088 -0.036 0.019
Interest Coverage Ratio -0.26 -0.23 -0.23 -0.24 -0.24 0.097 0.52 -0.21 1 0.09 0.5 -0.41 0.007
Price to Cash Flow per Share -0.36 -0.31 -0.29 -0.32 -0.33 0.068 0.47 0.25 0.09 1 0.14 -0.25 0.13
Return on Assets -0.26 -0.25 -0.23 -0.25 -0.25 0.18 0.34 -0.088 0.5 0.14 1 -0.35 0.26
Total Debt to Total Equity 0.82 0.81 0.79 0.81 0.81 -0.34 -0.34 -0.036 -0.41 -0.25 -0.35 1 -0.41
Weighted Average Cost of Capital -0.32 -0.32 -0.26 -0.28 -0.32 0.52 0.13 0.019 0.007 0.13 0.26 -0.41 1
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

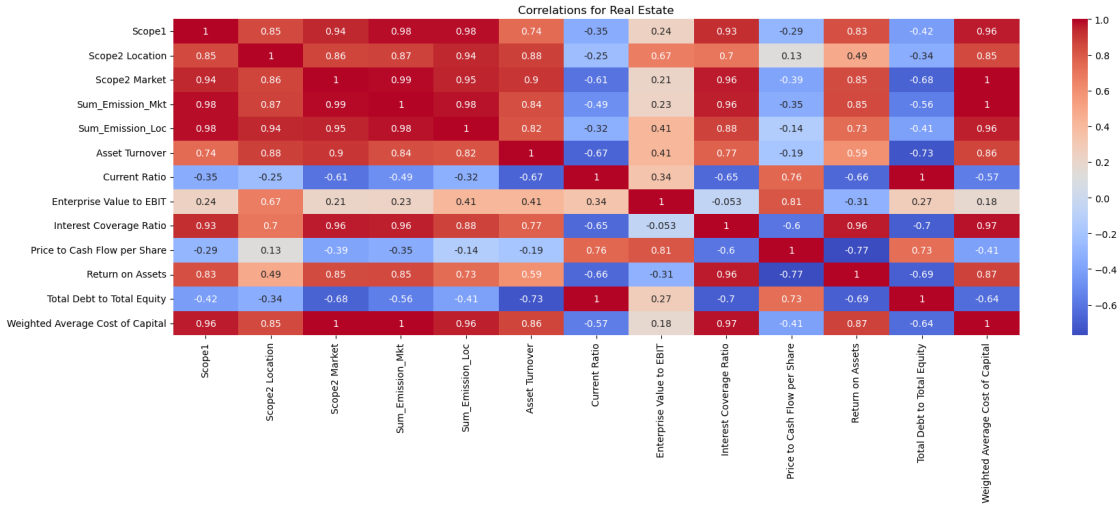

Correlations for Health Care
Scope1 1 0.87 0.83 0.96 0.97 -0.36 -0.14 0.074 -0.17 -0.19 -0.22 0.23 -0.0033
Scope2 Location 0.87 1 0.98 0.96 0.96 -0.053 -0.2 0.026 -0.22 -0.26 -0.31 0.33 0.031
Scope2 Market 0.83 0.98 1 0.95 0.93 0.017 -0.21 0.033 -0.19 -0.28 -0.3 0.3 -0.0092
Sum_Emission_Mkt 0.96 0.96 0.95 1 0.99 -0.19 -0.18 0.057 -0.19 -0.25 -0.27 0.27 -0.0064
Sum_Emission_Loc 0.97 0.96 0.93 0.99 1 -0.22 -0.18 0.053 -0.2 -0.23 -0.27 0.28 0.014
Asset Turnover -0.36 -0.053 0.017 -0.19 -0.22 1 -0.17 -0.27 -0.074 -0.31 0.27 -0.13 0.00026
Current Ratio -0.14 -0.2 -0.21 -0.18 -0.18 -0.17 1 -0.11 0.43 0.033 -0.087 -0.15 -0.21
Enterprise Value to EBIT 0.074 0.026 0.033 0.057 0.053 -0.27 -0.11 1 -0.055 0.17 -0.11 0.17 -0.049
Interest Coverage Ratio -0.17 -0.22 -0.19 -0.19 -0.2 -0.074 0.43 -0.055 1 0.23 0.024 -0.4 -0.21
Price to Cash Flow per Share -0.19 -0.26 -0.28 -0.25 -0.23 -0.31 0.033 0.17 0.23 1 -0.12 0.25 0.12
Return on Assets -0.22 -0.31 -0.3 -0.27 -0.27 0.27 -0.087 -0.11 0.024 -0.12 1 -0.4 0.31
Total Debt to Total Equity 0.23 0.33 0.3 0.27 0.28 -0.13 -0.15 0.17 -0.4 0.25 -0.4 1 0.053
Weighted Average Cost of Capital -0.0033 0.031 -0.0092 -0.0064 0.014 0.00026 -0.21 -0.049 -0.21 0.12 0.31 0.053 1
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

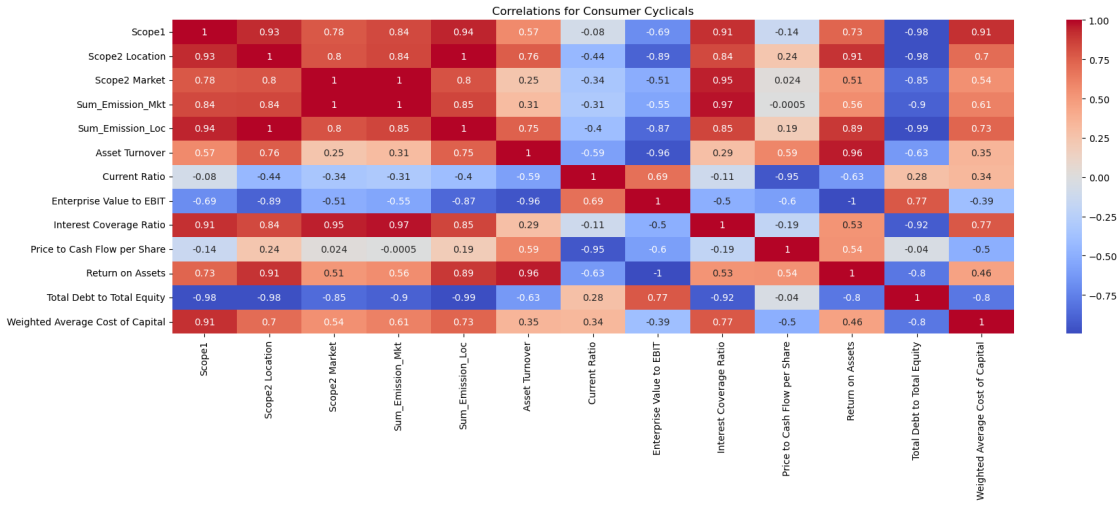

Correlations for Materials
Scope1 1 0.53 0.42 0.89 0.94 -0.12 -0.19 -0.093 -0.27 -0.59 -0.24 -0.2 0.49
Scope2 Location 0.53 1 0.94 0.83 0.8 0.079 0.23 -0.12 -0.081 -0.44 -0.31 -0.5 0.38
Scope2 Market 0.42 0.94 1 0.78 0.69 0.098 0.37 -0.1 -0.088 -0.4 -0.38 -0.48 0.4
Sum_Emission_Mkt 0.89 0.83 0.78 1 0.98 -0.031 0.052 -0.11 -0.23 -0.61 -0.36 -0.38 0.53
Sum_Emission_Loc 0.94 0.8 0.69 0.98 1 -0.05 -0.041 -0.12 -0.23 -0.61 -0.3 -0.35 0.51
Asset Turnover -0.12 0.079 0.098 -0.031 -0.05 1 0.047 -0.11 -0.0025 -0.27 -0.11 -0.35 0.062
Current Ratio -0.19 0.23 0.37 0.052 -0.041 0.047 1 -0.053 -0.12 0.1 -0.33 -0.24 -0.027
Enterprise Value to EBIT -0.093 -0.12 -0.1 -0.11 -0.12 -0.11 -0.053 1 -0.078 -0.036 -0.24 0.1 -0.24
Interest Coverage Ratio -0.27 -0.081 -0.088 -0.23 -0.23 -0.0025 -0.12 -0.078 1 0.21 0.46 -0.24 0.17
Price to Cash Flow per Share -0.59 -0.44 -0.4 -0.61 -0.61 -0.27 0.1 -0.036 0.21 1 0.64 0.23 -0.11
Return on Assets -0.24 -0.31 -0.38 -0.36 -0.3 -0.11 -0.33 -0.24 0.46 0.64 1 -0.027 0.16
Total Debt to Total Equity -0.2 -0.5 -0.48 -0.38 -0.35 -0.35 -0.24 0.1 -0.24 0.23 -0.027 1 -0.48
Weighted Average Cost of Capital 0.49 0.38 0.4 0.53 0.51 0.062 -0.027 -0.24 0.17 -0.11 0.16 -0.48 1
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6


Correlations for Communication Services
Scope1 1 0.41 0.66 0.75 0.45 0.11 -0.36 -0.32 -0.17 -0.58 -0.29 0.071 -0.23
Scope2 Location 0.41 1 0.57 0.57 1 0.3 0.25 -0.13 0.56 -0.11 0.31 -0.1 0.15
Scope2 Market 0.66 0.57 1 0.99 0.59 0.14 -0.0065 -0.09 0.24 -0.22 0.0097 -0.0087 0.066
Sum_Emission_Mkt 0.75 0.57 0.99 1 0.6 0.14 -0.06 -0.13 0.19 -0.28 -0.036 0.0031 0.024
Sum_Emission_Loc 0.45 1 0.59 0.6 1 0.3 0.22 -0.14 0.54 -0.14 0.29 -0.097 0.14
Asset Turnover 0.11 0.3 0.14 0.14 0.3 1 -0.018 -0.57 0.42 -0.34 0.53 -0.27 0.33
Current Ratio -0.36 0.25 -0.0065 -0.06 0.22 -0.018 1 0.6 0.64 0.7 0.44 -0.23 0.11
Enterprise Value to EBIT -0.32 -0.13 -0.09 -0.13 -0.14 -0.57 0.6 1 -0.032 0.87 -0.19 0.081 -0.18
Interest Coverage Ratio -0.17 0.56 0.24 0.19 0.54 0.42 0.64 -0.032 1 0.23 0.82 -0.31 0.46
Price to Cash Flow per Share -0.58 -0.11 -0.22 -0.28 -0.14 -0.34 0.7 0.87 0.23 1 0.16 -0.091 0.14
Return on Assets -0.29 0.31 0.0097 -0.036 0.29 0.53 0.44 -0.19 0.82 0.16 1 -0.38 0.46
Total Debt to Total Equity 0.071 -0.1 -0.0087 0.0031 -0.097 -0.27 -0.23 0.081 -0.31 -0.091 -0.38 1 -0.32
Weighted Average Cost of Capital -0.23 0.15 0.066 0.024 0.14 0.33 0.11 -0.18 0.46 0.14 0.46 -0.32 1
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4


Correlations for Utilities
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

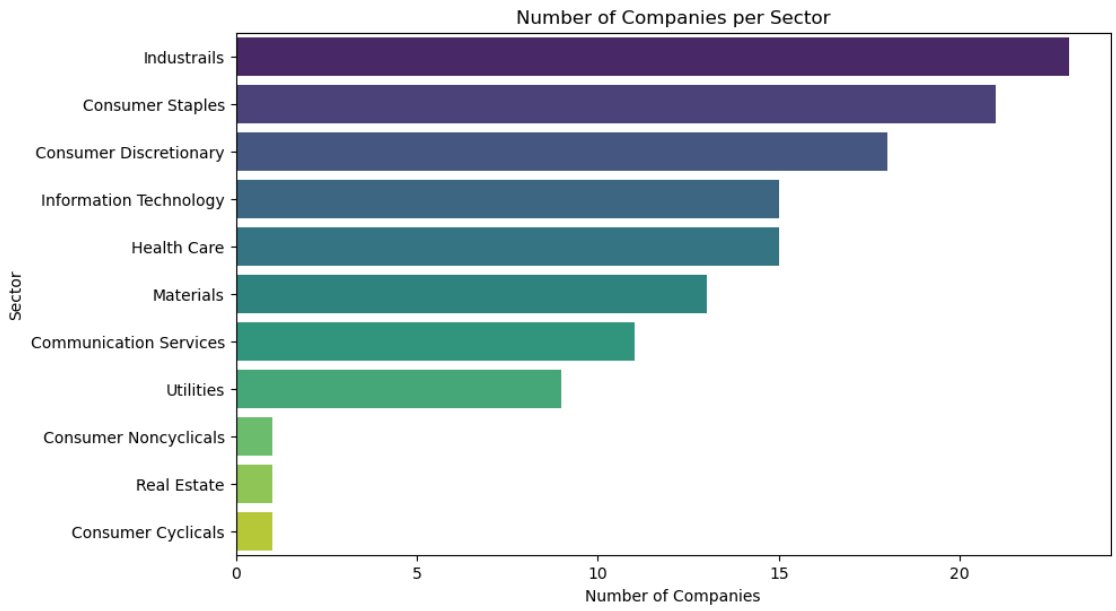

Correlations for Consumer Noncyclicals
Scope1
Scope2 Location
Scope2 Market
Sum_Emission_Mkt
Sum_Emission_Loc
Asset Turnover
Current Ratio
Enterprise Value to EBIT
Interest Coverage Ratio
Price to Cash Flow per Share
Return on Assets
Total Debt to Total Equity
Weighted Average Cost of Capital
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75





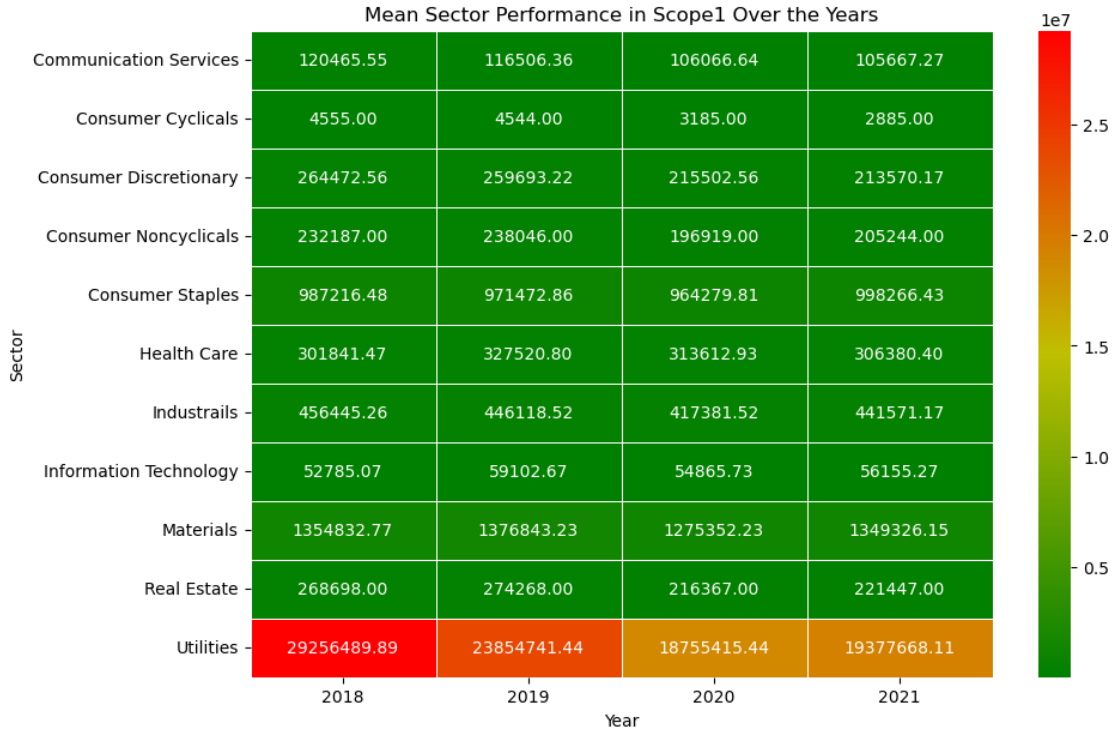
Company Distribution Across Sectors

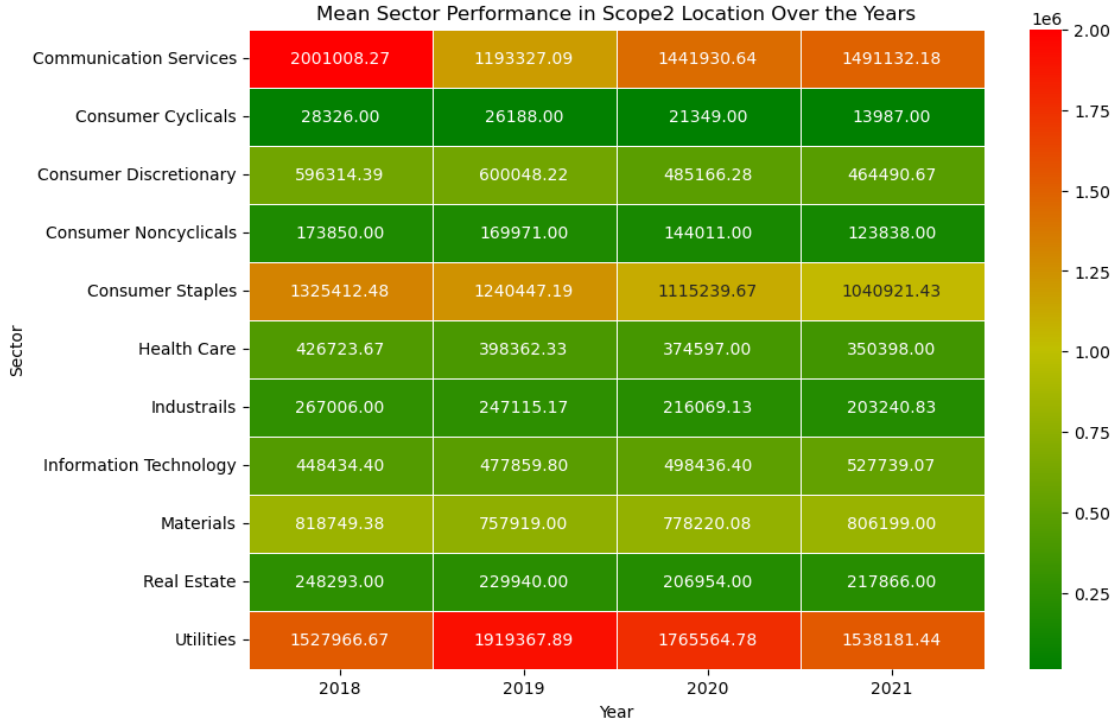
[40]:
```
# Count the number of unique companies per sector
sectors = final_data[final_data.index.get_level_values('Year')==2018]
company_count_per_sector = sectors['Sector'].value_counts()
plt.figure(figsize=(10,6))
sns.barplot(x=company_count_per_sector.values, y=company_count_per_sector.
 ↪index, palette='viridis')
plt.title('Number of Companies per Sector')
plt.xlabel('Number of Companies')
plt.ylabel('Sector')
plt.show()
```

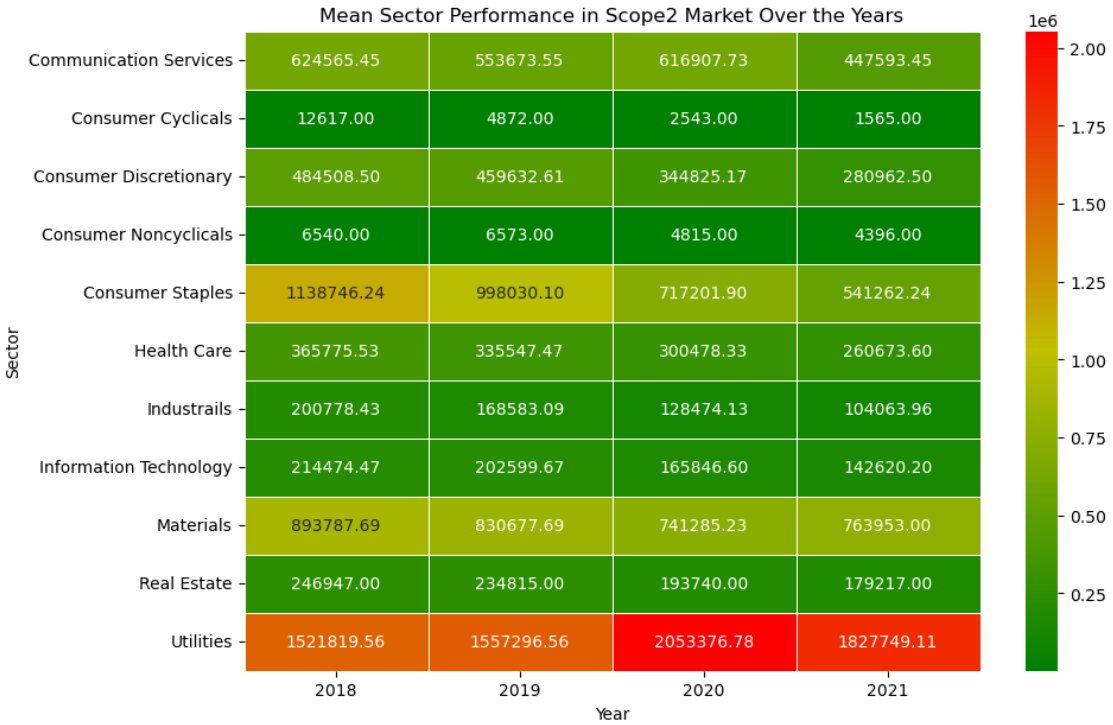


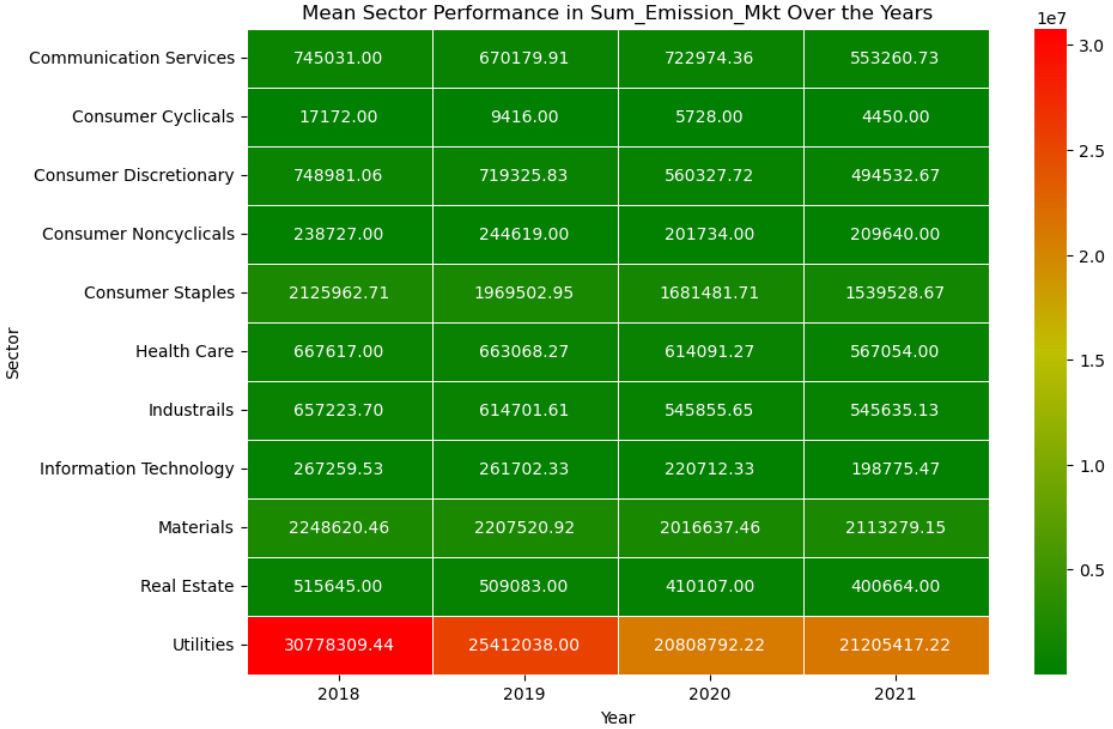
Heatmap Sector Performance Visualise sector performance over time using heatmaps

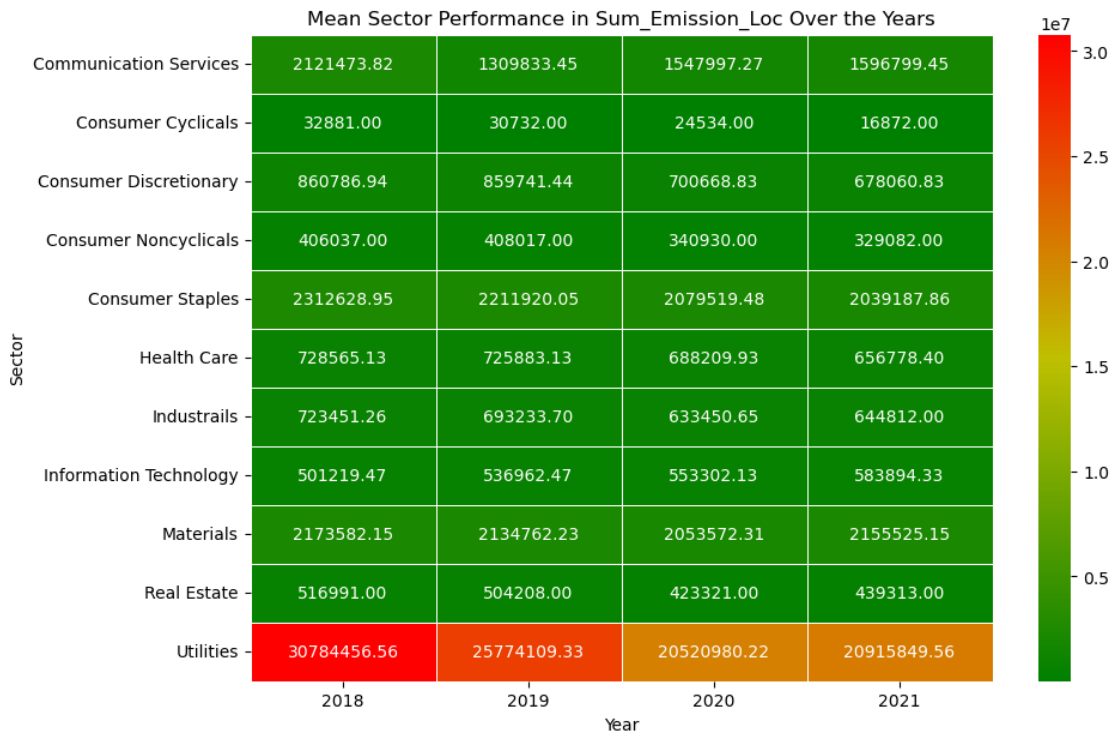
[41]:
```
def plot_sector_performance(df, metrics):
    """
    Plots the sector performance over the years for the given metrics using a␣
 ↪heatmap

    Parameters:
        df (pandas.DataFrame): The DataFrame containing sector and year␣
 ↪information
        metrics (list): The list of metrics to plot

    Returns:
        None
    """
    df_sector_mean = df.groupby(['Year', 'Sector'])[metrics].mean()

    for metric in metrics:
        pivot_df = df_sector_mean[f'{metric}'].unstack(level=0)
        plt.figure(figsize=(10,7))
        cmap = LinearSegmentedColormap.from_list('gr',["g", "y", "r"], N=256)
        sns.heatmap(pivot_df, annot=True, fmt=".2f", cmap=cmap, linewidths=0.5)
        plt.title(f'Mean Sector Performance in {metric} Over the Years')
        plt.show()

plot_sector_performance(final_data, final_data.columns[1:])
```

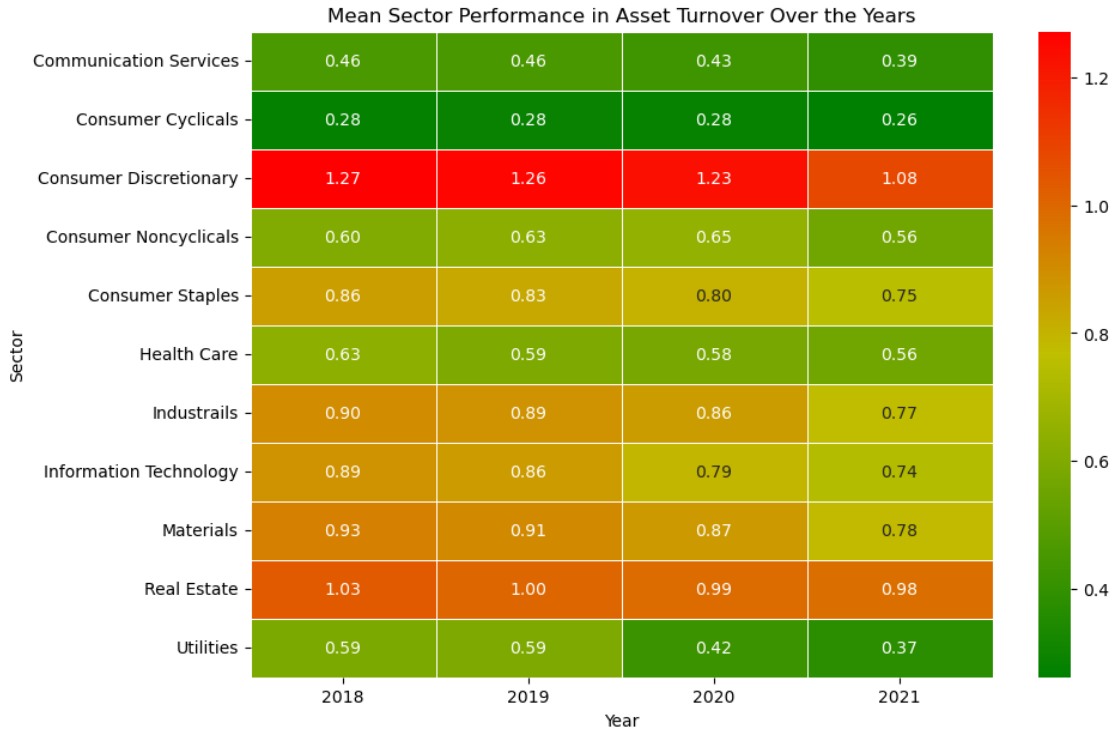

Mean Sector Performance in Scope1 Over the Years
Sector
Year
1e7
Communication Services 120465.55 116506.36 106066.64 105667.27
Consumer Cyclicals 4555.00 4544.00 3185.00 2885.00
Consumer Discretionary 264472.56 259693.22 215502.56 213570.17
Consumer Noncyclicals 232187.00 238046.00 196919.00 205244.00
Consumer Staples 987216.48 971472.86 964279.81 998266.43
Health Care 301841.47 327520.80 313612.93 306380.40
Industrails 456445.26 446118.52 417381.52 441571.17
Information Technology 52785.07 59102.67 54865.73 56155.27
Materials 1354832.77 1376843.23 1275352.23 1349326.15
Real Estate 268698.00 274268.00 216367.00 221447.00
Utilities 29256489.89 23854741.44 18755415.44 19377668.11
2018 2019 2020 2021
0.5 1.0 1.5 2.0 2.5

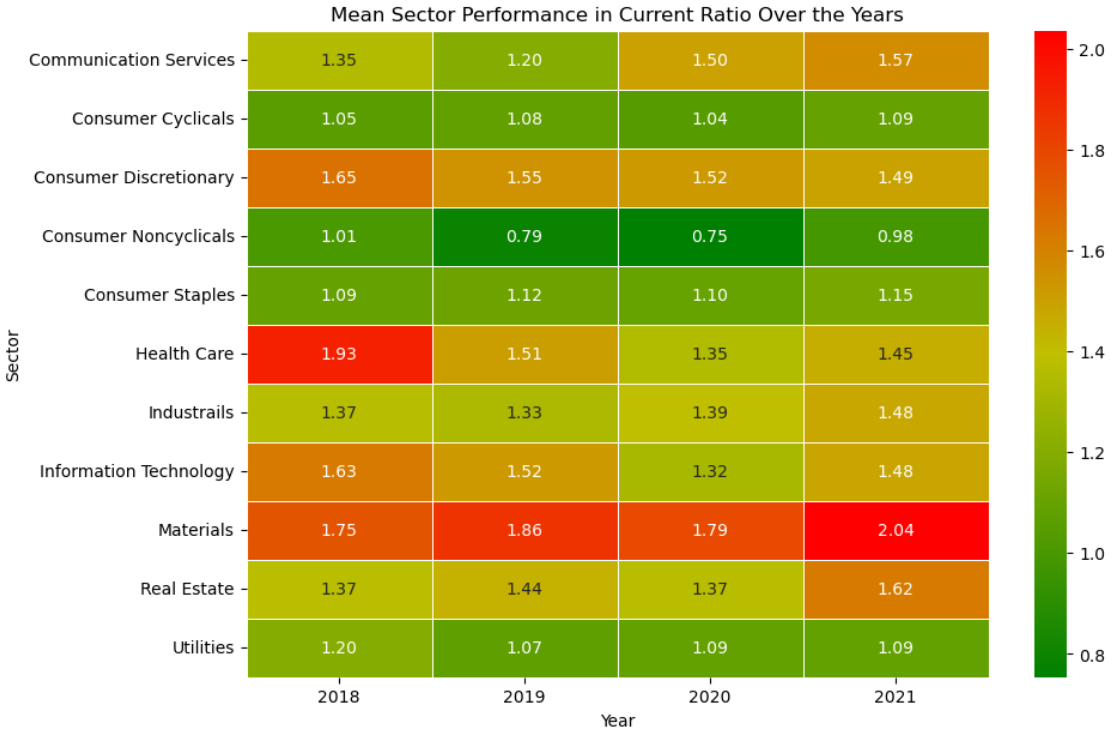

Mean Sector Performance in Scope2 Location Over the Years
Sector
Year
1e6
Communication Services 2001008.27 1193327.09 1441930.64 1491132.18
Consumer Cyclicals 28326.00 26188.00 21349.00 13987.00
Consumer Discretionary 596314.39 600048.22 485166.28 464490.67
Consumer Noncyclicals 173850.00 169971.00 144011.00 123838.00
Consumer Staples 1325412.48 1240447.19 1115239.67 1040921.43
Health Care 426723.67 398362.33 374597.00 350398.00
Industrails 267006.00 247115.17 216069.13 203240.83
Information Technology 448434.40 477859.80 498436.40 527739.07
Materials 818749.38 757919.00 778220.08 806199.00
Real Estate 248293.00 229940.00 206954.00 217866.00
Utilities 1527966.67 1919367.89 1765564.78 1538181.44
2018 2019 2020 2021
0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00


Mean Sector Performance in Scope2 Market Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
624565.45
553673.55
616907.73
447593.45
12617.00
4872.00
2543.00
1565.00
484508.50
459632.61
344825.17
280962.50
6540.00
6573.00
4815.00
4396.00
1138746.24
998030.10
717201.90
541262.24
365775.53
335547.47
300478.33
260673.60
200778.43
168583.09
128474.13
104063.96
214474.47
202599.67
165846.60
142620.20
893787.69
830677.69
741285.23
763953.00
246947.00
234815.00
193740.00
179217.00
1521819.56
1557296.56
2053376.78
1827749.11
2018
2019
2020
2021
Year
1e6
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25

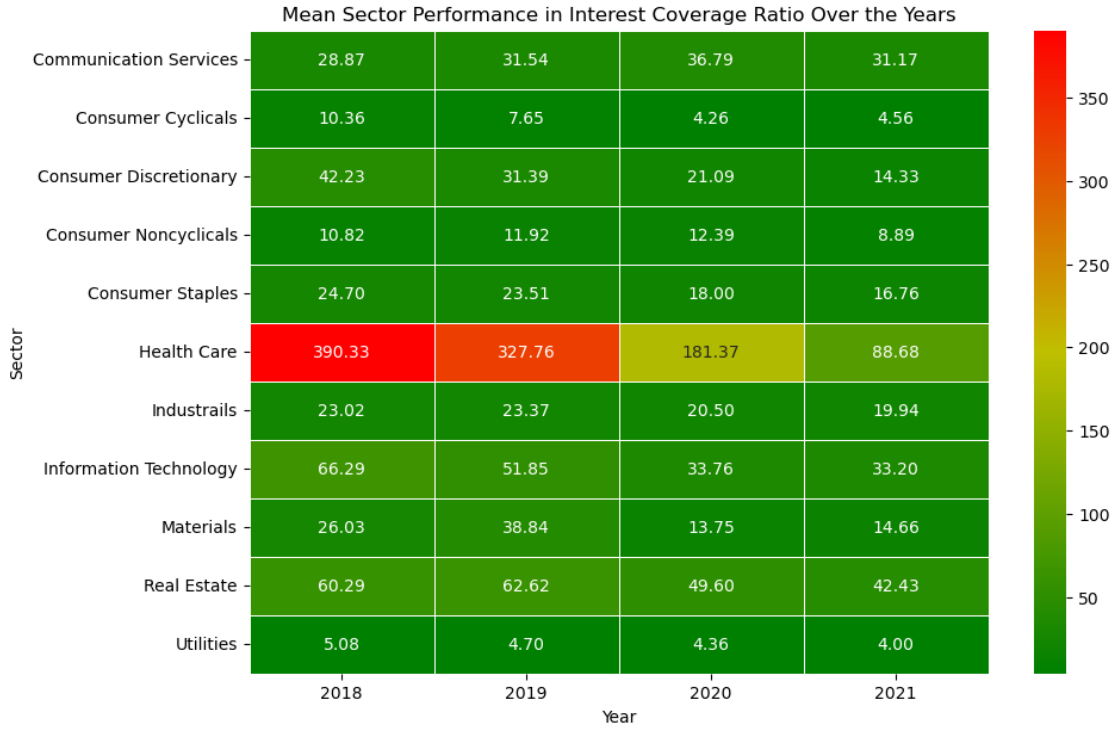

Mean Sector Performance in Sum_Emission_Mkt Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
745031.00
670179.91
722974.36
553260.73
17172.00
9416.00
5728.00
4450.00
748981.06
719325.83
560327.72
494532.67
238727.00
244619.00
201734.00
209640.00
2125962.71
1969502.95
1681481.71
1539528.67
667617.00
663068.27
614091.27
567054.00
657223.70
614701.61
545855.65
545635.13
267259.53
261702.33
220712.33
198775.47
2248620.46
2207520.92
2016637.46
2113279.15
515645.00
509083.00
410107.00
400664.00
30778309.44
25412038.00
20808792.22
21205417.22
2018
2019
2020
2021
Year
1e7
3.0
2.5
2.0
1.5
1.0
0.5

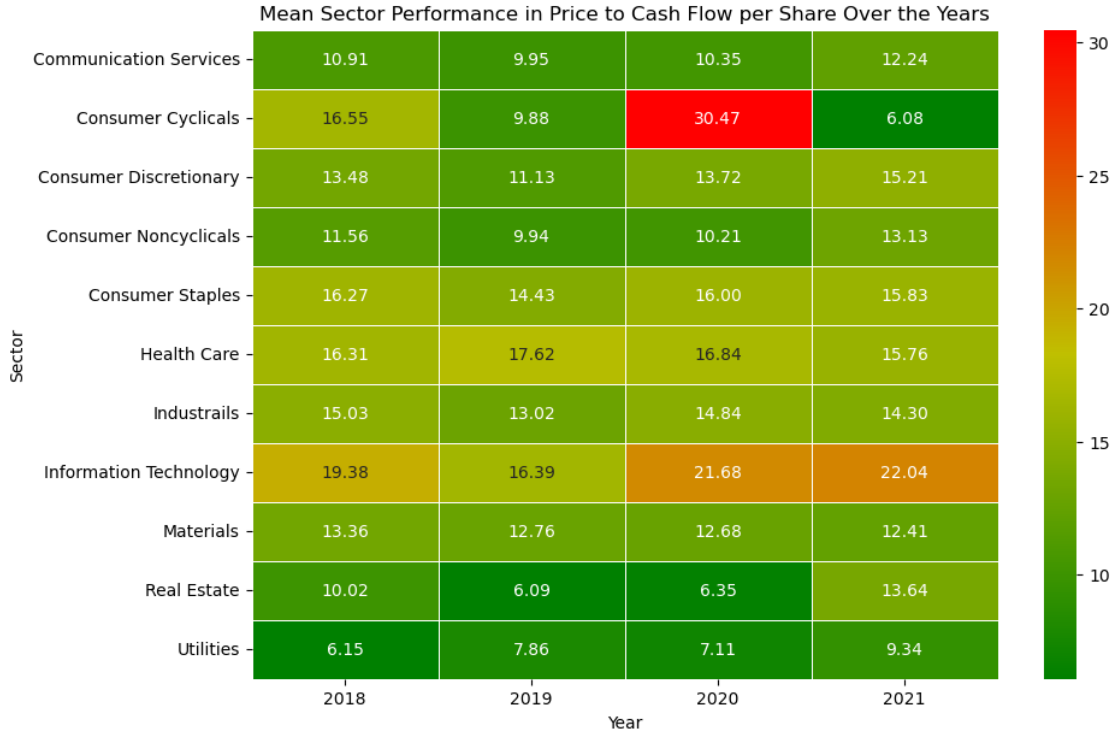

Mean Sector Performance in Sum_Emission_Loc Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
2121473.82
1309833.45
1547997.27
1596799.45
32881.00
30732.00
24534.00
16872.00
860786.94
859741.44
700668.83
678060.83
406037.00
408017.00
340930.00
329082.00
2312628.95
2211920.05
2079519.48
2039187.86
728565.13
725883.13
688209.93
656778.40
723451.26
693233.70
633450.65
644812.00
501219.47
536962.47
553302.13
583894.33
2173582.15
2134762.23
2053572.31
2155525.15
516991.00
504208.00
423321.00
439313.00
30784456.56
25774109.33
20520980.22
20915849.56
2018
2019
2020
2021
Year
1e7
3.0
2.5
2.0
1.5
1.0
0.5

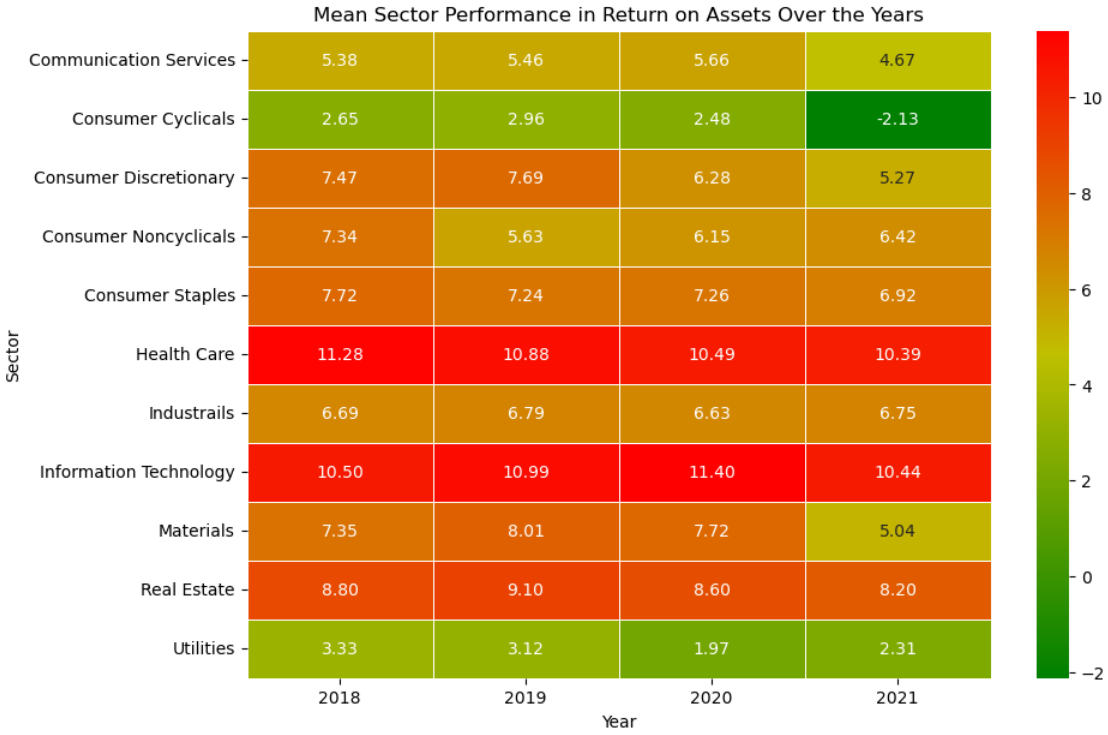

Mean Sector Performance in Asset Turnover Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
0.46
0.46
0.43
0.39
0.28
0.28
0.28
0.26
1.27
1.26
1.23
1.08
0.60
0.63
0.65
0.56
0.86
0.83
0.80
0.75
0.63
0.59
0.58
0.56
0.90
0.89
0.86
0.77
0.89
0.86
0.79
0.74
0.93
0.91
0.87
0.78
1.03
1.00
0.99
0.98
0.59
0.59
0.42
0.37
2018
2019
2020
2021
Year
1.2
1.0
0.8
0.6
0.4

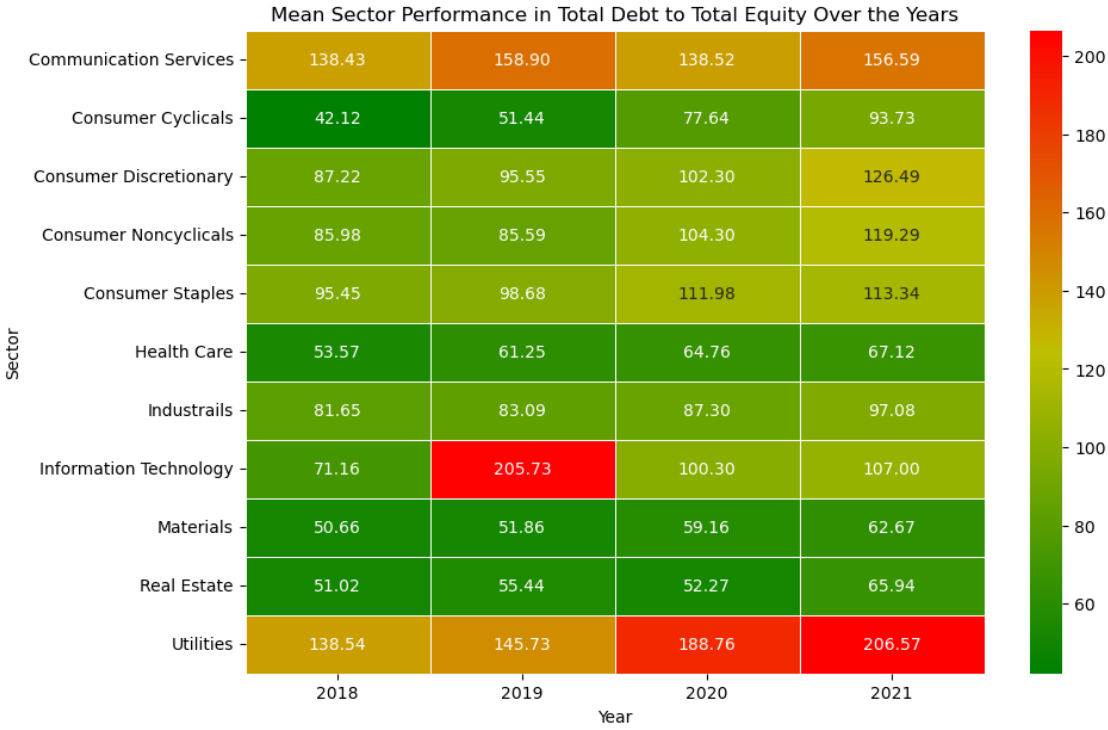

Mean Sector Performance in Current Ratio Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
1.35 1.20 1.50 1.57
1.05 1.08 1.04 1.09
1.65 1.55 1.52 1.49
1.01 0.79 0.75 0.98
1.09 1.12 1.10 1.15
1.93 1.51 1.35 1.45
1.37 1.33 1.39 1.48
1.63 1.52 1.32 1.48
1.75 1.86 1.79 2.04
1.37 1.44 1.37 1.62
1.20 1.07 1.09 1.09
2018 2019 2020 2021
Year
2.0
1.8
1.6
1.4
1.2
1.0
0.8

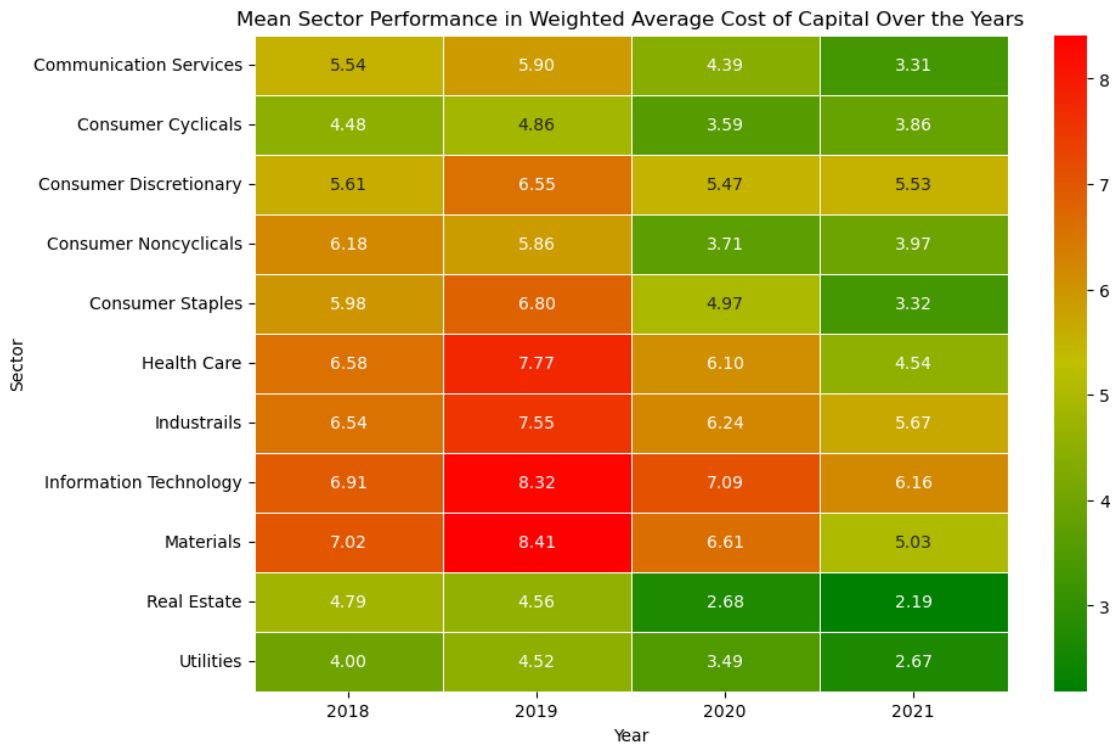
Mean Sector Performance in Enterprise Value to EBIT Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
19.72 20.52 28.25 35.23
16.07 15.12 18.10 937.94
66.80 19.47 19.08 77.59
18.15 16.51 16.19 19.10
18.47 16.26 17.48 18.12
18.38 18.96 33.91 21.15
16.89 14.15 16.70 19.93
32.12 27.05 52.11 71.50
18.69 13.20 20.08 42.89
9.89 8.20 7.75 9.48
14.08 21.91 16.11 21.66
2018 2019 2020 2021
Year
800
600
400
200


Mean Sector Performance in Interest Coverage Ratio Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
28.87 31.54 36.79 31.17
10.36 7.65 4.26 4.56
42.23 31.39 21.09 14.33
10.82 11.92 12.39 8.89
24.70 23.51 18.00 16.76
390.33 327.76 181.37 88.68
23.02 23.37 20.50 19.94
66.29 51.85 33.76 33.20
26.03 38.84 13.75 14.66
60.29 62.62 49.60 42.43
5.08 4.70 4.36 4.00
2018 2019 2020 2021
Year
350
300
250
200
150
100
50


Mean Sector Performance in Price to Cash Flow per Share Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
10.91 9.95 10.35 12.24
16.55 9.88 30.47 6.08
13.48 11.13 13.72 15.21
11.56 9.94 10.21 13.13
16.27 14.43 16.00 15.83
16.31 17.62 16.84 15.76
15.03 13.02 14.84 14.30
19.38 16.39 21.68 22.04
13.36 12.76 12.68 12.41
10.02 6.09 6.35 13.64
6.15 7.86 7.11 9.34
2018 2019 2020 2021
Year
30
25
20
15
10


Mean Sector Performance in Return on Assets Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
5.38
5.46
5.66
4.67
2.65
2.96
2.48
-2.13
7.47
7.69
6.28
5.27
7.34
5.63
6.15
6.42
7.72
7.24
7.26
6.92
11.28
10.88
10.49
10.39
6.69
6.79
6.63
6.75
10.50
10.99
11.40
10.44
7.35
8.01
7.72
5.04
8.80
9.10
8.60
8.20
3.33
3.12
1.97
2.31
2018
2019
2020
2021
Year
10
8
6
4
2
0
−2

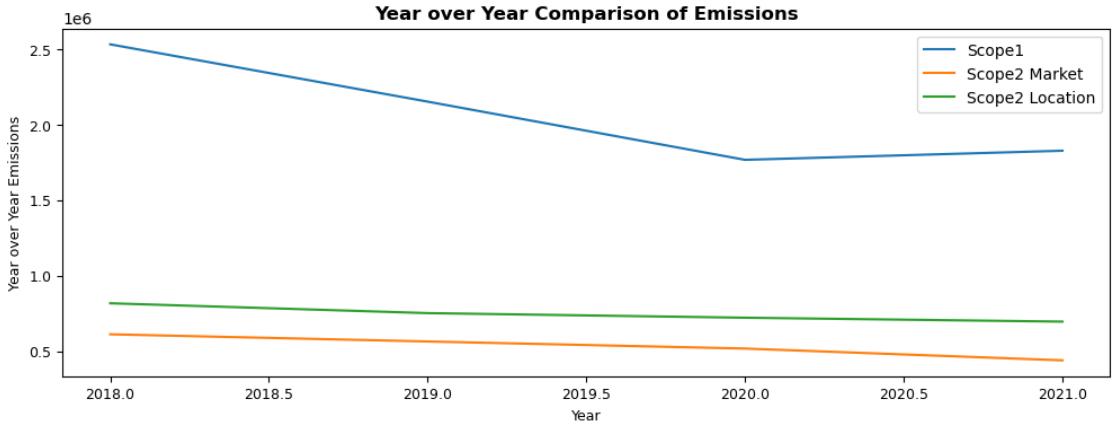

Mean Sector Performance in Total Debt to Total Equity Over the Years
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector
138.43
158.90
138.52
156.59
42.12
51.44
77.64
93.73
87.22
95.55
102.30
126.49
85.98
85.59
104.30
119.29
95.45
98.68
111.98
113.34
53.57
61.25
64.76
67.12
81.65
83.09
87.30
97.08
71.16
205.73
100.30
107.00
50.66
51.86
59.16
62.67
51.02
55.44
52.27
65.94
138.54
145.73
188.76
206.57
2018
2019
2020
2021
Year
200
180
160
140
120
100
80
60



[42]:
```
line1 = final_data.groupby(['Year'])['Scope1','Scope2 Market','Scope2␣
↪Location'].mean()
line1.plot(figsize=(12,4),fontsize=9)
plt.ylabel('Year over Year Emissions', fontsize = 9 )
plt.xlabel('Year', fontsize = 9)
plt.title('Year over Year Comparison of Emissions',fontsize =␣
↪12,fontweight='bold')
```

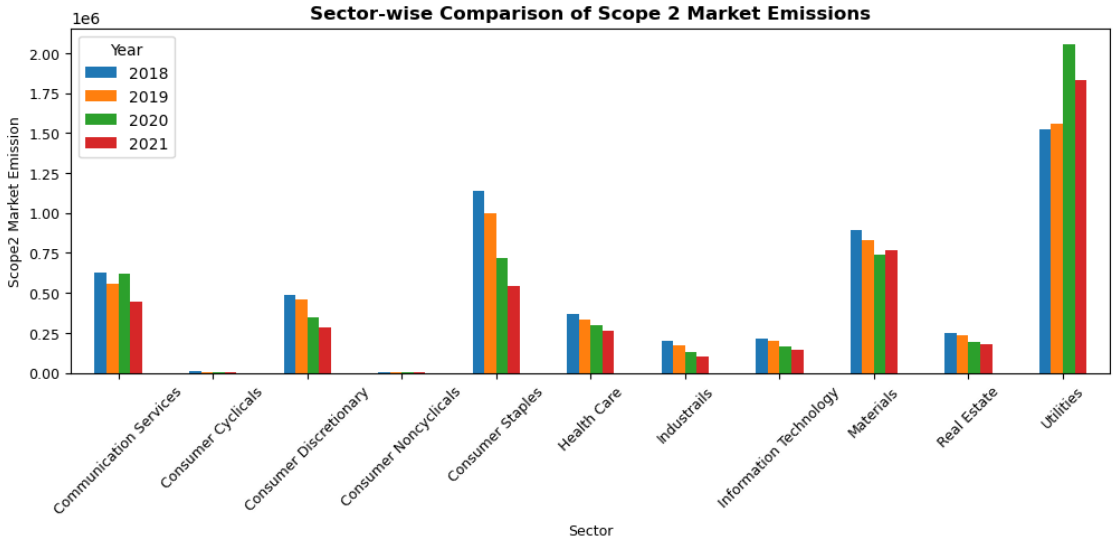
```
/var/folders/9p/xr85g4y15w1fy3_389l3s_nh0000gn/T/ipykernel_86452/3289259431.py:1
: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of
keys) will be deprecated, use a list instead.
  line1 = final_data.groupby(['Year'])['Scope1','Scope2 Market','Scope2
Location'].mean()
```

[42]: Text(0.5, 1.0, 'Year over Year Comparison of Emissions')



[43]:
```
bar1 = final_data.groupby(['Sector','Year'])['Scope2 Market'].mean().
↪unstack(level=1)
bar1.plot(kind='bar',figsize=(12,4),fontsize=9)
plt.ylabel('Scope2 Market Emission', fontsize = 9 )
plt.xlabel('Sector', fontsize = 9)
plt.xticks(rotation = 45)
plt.title('Sector-wise Comparison of Scope 2 Market Emissions',fontsize =␣
↪12,fontweight='bold')
```

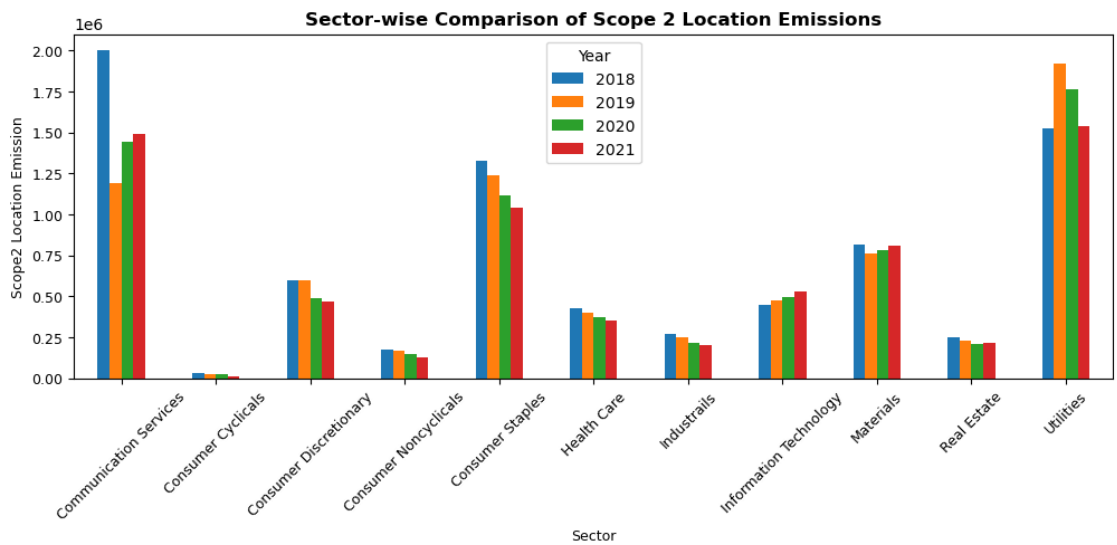
[43]: Text(0.5, 1.0, 'Sector-wise Comparison of Scope 2 Market Emissions')



[44]:
```
bar2 = final_data.groupby(['Sector','Year'])['Scope2 Location'].mean().
↪unstack(level=1)
```

```
bar2.plot(kind='bar',figsize=(12,4),fontsize=9)
plt.ylabel('Scope2 Location Emission', fontsize = 9 )
plt.xlabel('Sector', fontsize = 9)
plt.xticks(rotation =45)
plt.title('Sector-wise Comparison of Scope 2 Location Emissions',fontsize =␣
↪12,fontweight='bold')
```

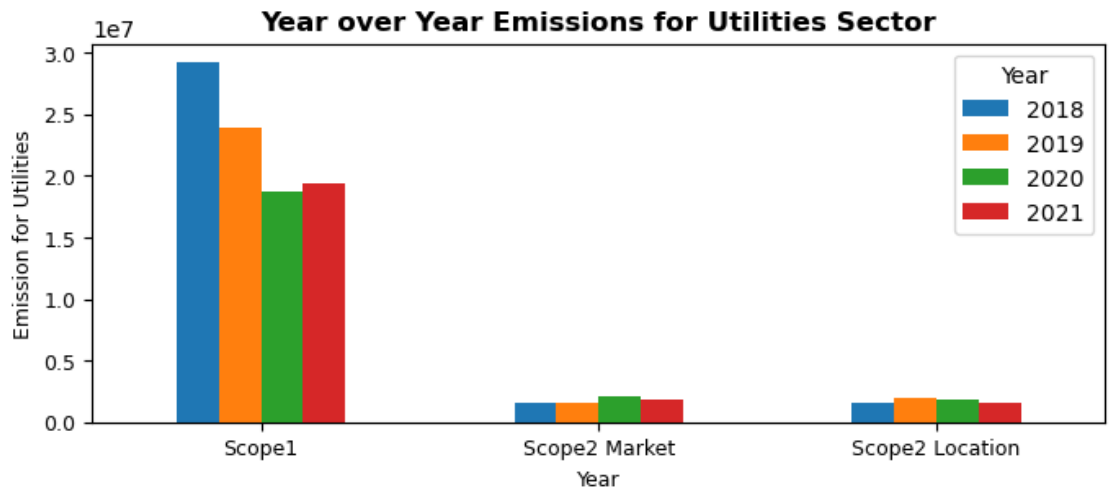
[44]: Text(0.5, 1.0, 'Sector-wise Comparison of Scope 2 Location Emissions')



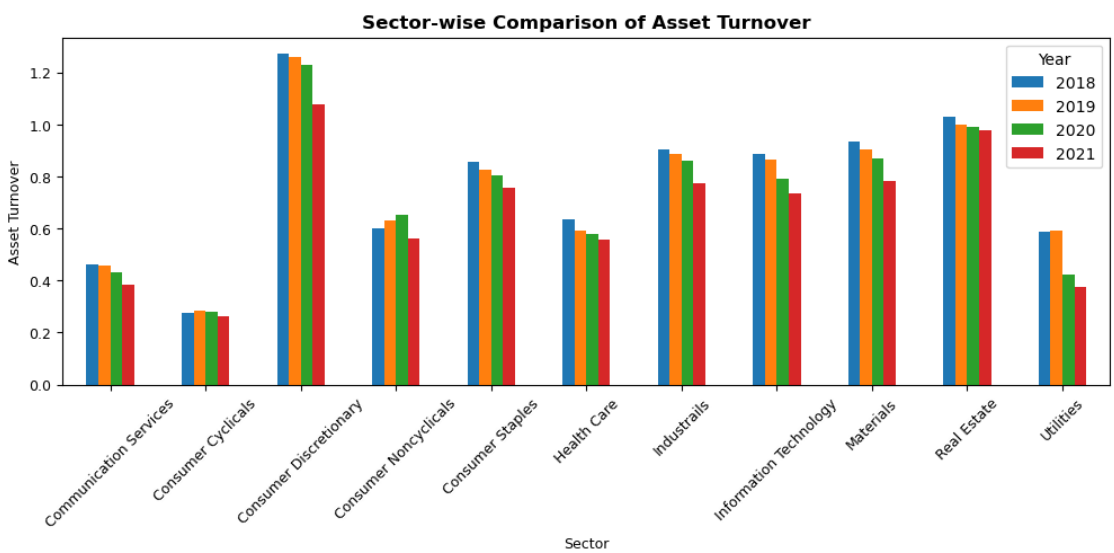
[45]:
```
bar3 = final_data.groupby(['Sector','Year'])['Scope1','Scope2 Market','Scope2␣
↪Location'].mean().unstack(level=1).T
bar3 = bar3['Utilities'].unstack(level=-1)
bar3.plot(kind='bar',figsize=(8,3),fontsize=9)
plt.ylabel('Emission for Utilities', fontsize = 9 )
plt.xlabel('Year', fontsize = 9)
plt.xticks(rotation = 0)
plt.title('Year over Year Emissions for Utilities Sector',fontsize =␣
↪12,fontweight='bold')
```

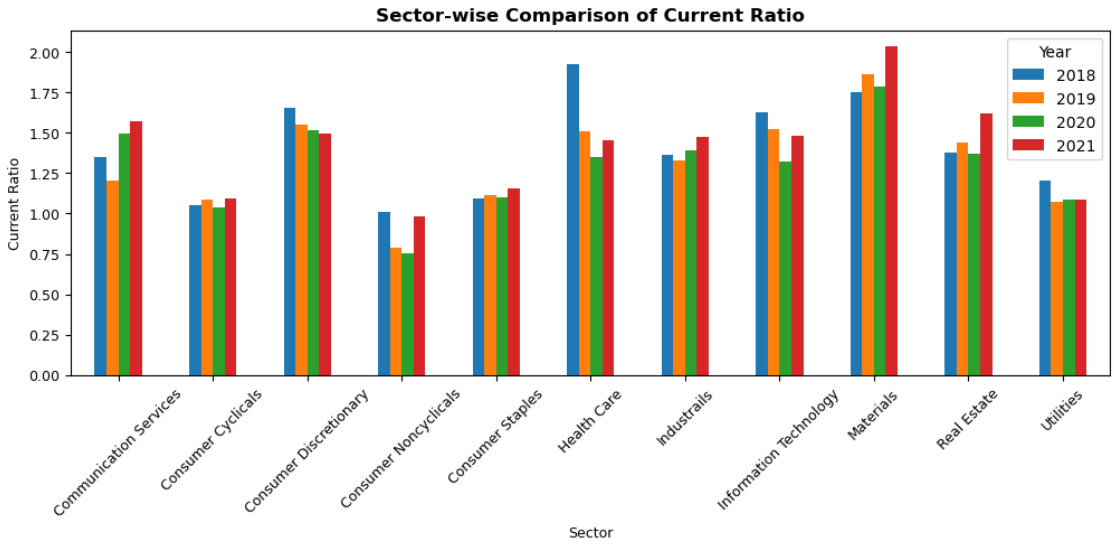
```
/var/folders/9p/xr85g4y15w1fy3_389l3s_nh0000gn/T/ipykernel_86452/4070084277.py:1
: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of
keys) will be deprecated, use a list instead.
  bar3 = final_data.groupby(['Sector','Year'])['Scope1','Scope2 Market','Scope2
Location'].mean().unstack(level=1).T
```

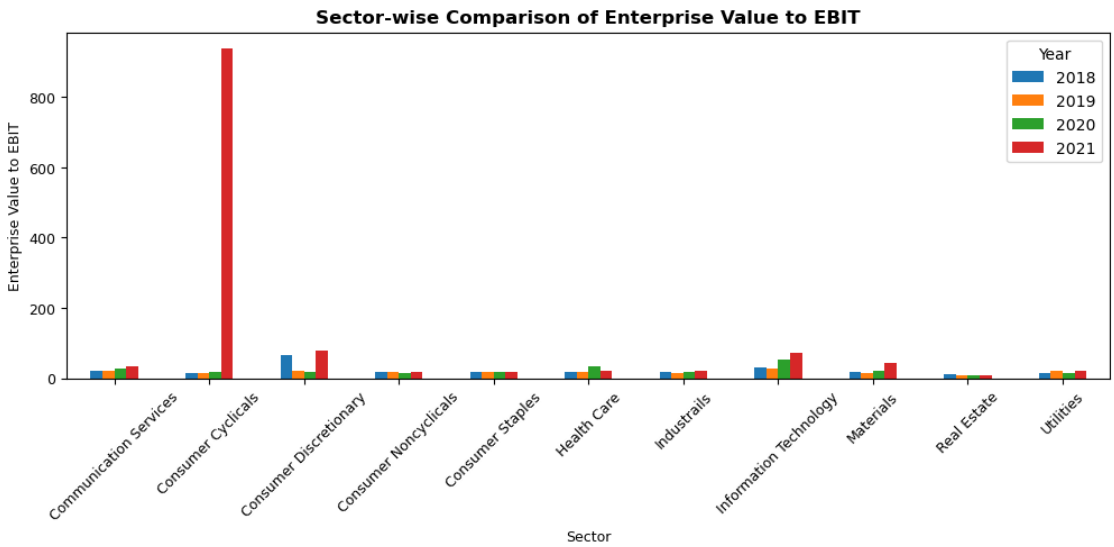
[45]: Text(0.5, 1.0, 'Year over Year Emissions for Utilities Sector')

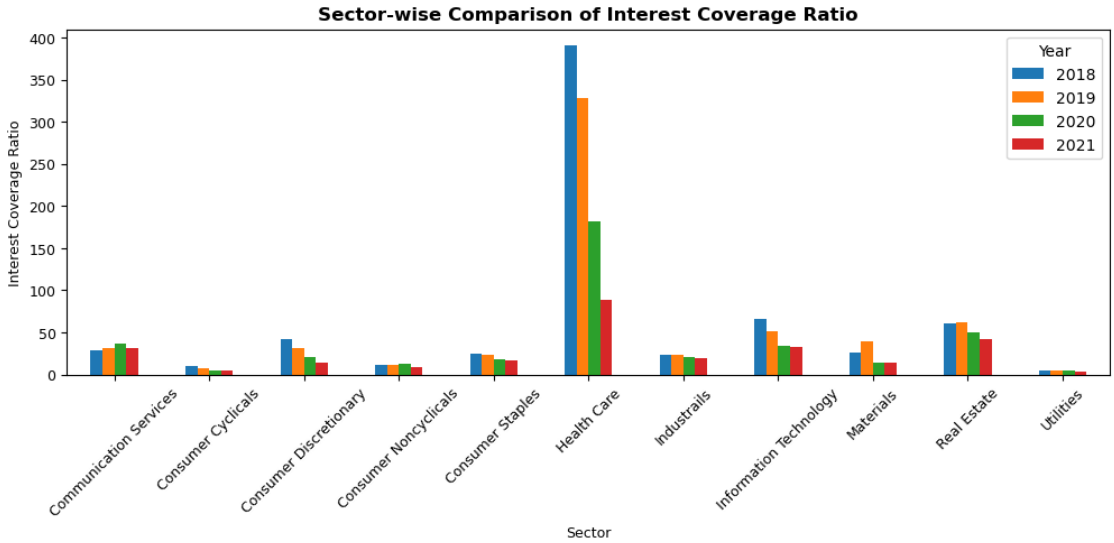


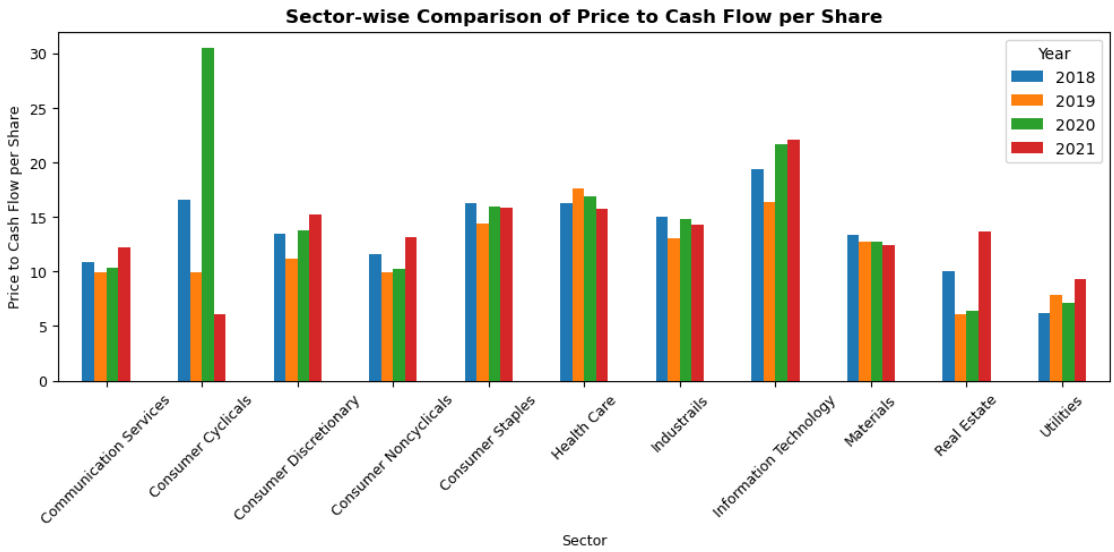
[46]:
```
fin_metrics = ['Asset Turnover', 'Current Ratio', 'Enterprise Value to EBIT',␣
↪'Interest Coverage Ratio',
               'Price to Cash Flow per Share', 'Return on Assets' , 'Total Debt␣
↪to Total Equity',
               'Weighted Average Cost of Capital']
for i in fin_metrics:
    bar4 = final_data.groupby(['Sector','Year'])[i].mean().unstack(level=1)
    bar4.plot(kind='bar',figsize=(12,4),fontsize=9)
    plt.ylabel(i, fontsize = 9 )
    plt.xlabel('Sector', fontsize = 9)
    plt.xticks(rotation =45)
    plt.title(f'Sector-wise Comparison of {i}' ,fontsize = 12,fontweight='bold')
```

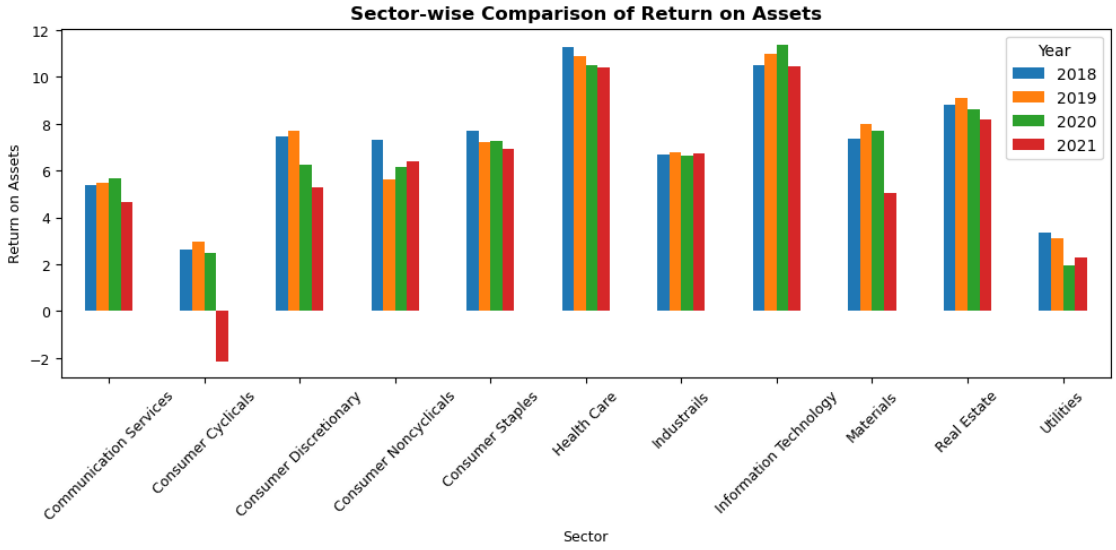


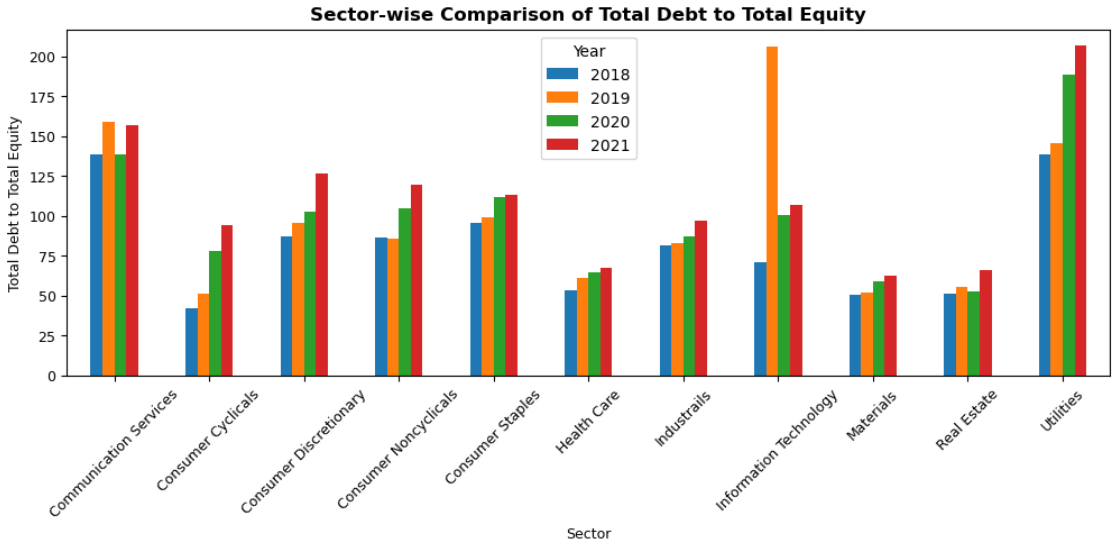

Sector-wise Comparison of Current Ratio
Current Ratio
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector

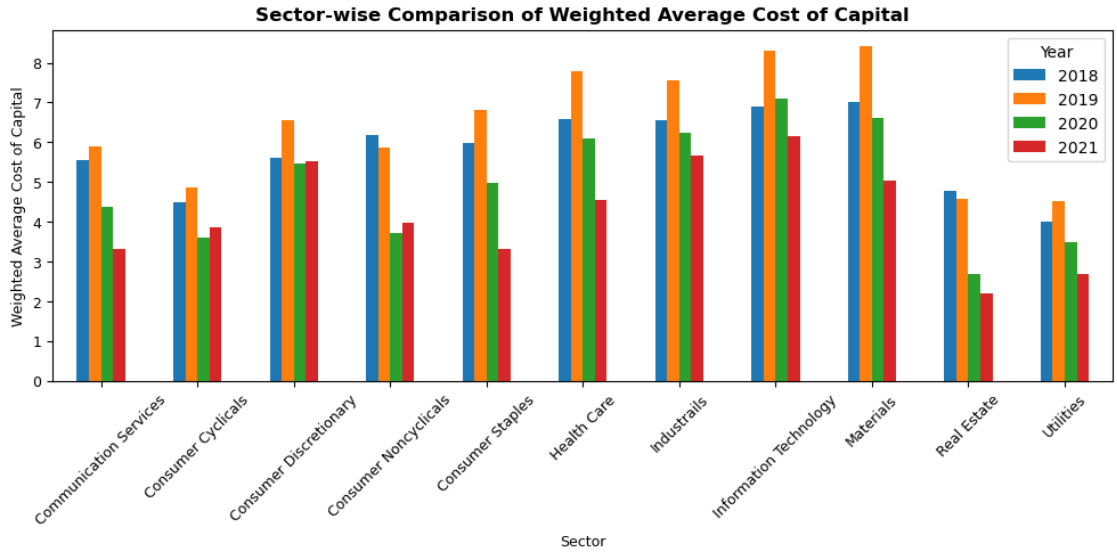

Sector-wise Comparison of Enterprise Value to EBIT
Enterprise Value to EBIT
800
600
400
200
0
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities
Sector


Sector-wise Comparison of Interest Coverage Ratio
Interest Coverage Ratio
Sector
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities


Sector-wise Comparison of Price to Cash Flow per Share
Price to Cash Flow per Share
Sector
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities


Sector-wise Comparison of Return on Assets
Return on Assets
Sector
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities


Sector-wise Comparison of Total Debt to Total Equity
Total Debt to Total Equity
Sector
Year
2018
2019
2020
2021
Communication Services
Consumer Cyclicals
Consumer Discretionary
Consumer Noncyclicals
Consumer Staples
Health Care
Industrails
Information Technology
Materials
Real Estate
Utilities



2.2 Regression

[47]:
```
final_data.columns
```

[47]:
```
Index(['Sector', 'Scope1', 'Scope2 Location', 'Scope2 Market',
       'Sum_Emission_Mkt', 'Sum_Emission_Loc', 'Asset Turnover',
       'Current Ratio', 'Enterprise Value to EBIT', 'Interest Coverage Ratio',
       'Price to Cash Flow per Share', 'Return on Assets',
       'Total Debt to Total Equity', 'Weighted Average Cost of Capital'],
      dtype='object')
```

[48]:
```
# create a list of variables that you want to use as Xs
x_vars = ['Asset Turnover','Current Ratio', 'Enterprise Value to EBIT',␣
 ↪'Interest Coverage Ratio',
          'Price to Cash Flow per Share', 'Return on Assets', 'Total Debt to␣
 ↪Total Equity',
          'Weighted Average Cost of Capital']
# add constant to Xs df
X = sm.add_constant(final_data[x_vars])
```

2.2.1 METHOD 1

Model 1 – Scope1

[49]:
```
# create Y df
Y1 = np.log(final_data[['Scope1']])

# specify the model
mod1 = PanelOLS(Y1, X)
```

```python
res1 = mod1.fit()
print(res1)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                 Scope1   R-squared:                        0.1787
Estimator:                   PanelOLS   R-squared (Between):              0.2161
No. Observations:                 512   R-squared (Within):              -4.2813
Date:                Thu, Jun 15 2023   R-squared (Overall):              0.1787
Time:                        13:58:42   Log-likelihood                   -1122.4
Cov. Estimator:            Unadjusted
                                        F-statistic:                      13.678
Entities:                         128   P-value                           0.0000
Avg Obs:                       4.0000   Distribution:                   F(8,503)
Min Obs:                       4.0000
Max Obs:                       4.0000   F-statistic (robust):             13.678
                                        P-value                           0.0000
Time periods:                       4   Distribution:                   F(8,503)
Avg Obs:                      128.00
Min Obs:                      128.00
Max Obs:                      128.00

                                        Parameter Estimates
================================================================================
====================
                                 Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
const                               14.487     0.3999     36.230     0.0000
13.701      15.273
Asset Turnover                     -0.6201     0.1972    -3.1440     0.0018
-1.0075     -0.2326
Current Ratio                      -0.1906     0.1487    -1.2819     0.2005
-0.4826      0.1015
Enterprise Value to EBIT           -0.0048     0.0014    -3.4329     0.0006
-0.0075     -0.0020
Interest Coverage Ratio            -0.0005     0.0003    -1.7226     0.0856
-0.0012   7.727e-05
Price to Cash Flow per Share       -0.0561     0.0118    -4.7436     0.0000
-0.0794     -0.0329
Return on Assets                   -0.0648     0.0213    -3.0393     0.0025
-0.1067     -0.0229
Total Debt to Total Equity          0.0001     0.0008     0.1323     0.8948
-0.0015      0.0017
Weighted Average Cost of Capital   -0.0717     0.0511    -1.4017     0.1616
-0.1721      0.0288
```

```
==============================================================================
====================
```

Model 2 – Scope2 Location

[50]:
```python
# create Y df
Y2 = np.log(final_data[['Scope2 Location']])

# specify the model
mod2 = PanelOLS(Y2, X)
res2 = mod2.fit()
print(res2)
```

```
                          PanelOLS Estimation Summary
==============================================================================
Dep. Variable:        Scope2 Location   R-squared:                       0.1083
Estimator:                   PanelOLS   R-squared (Between):             0.1348
No. Observations:                 512   R-squared (Within):             -1.2580
Date:                Thu, Jun 15 2023   R-squared (Overall):             0.1083
Time:                        13:58:42   Log-likelihood                  -958.23
Cov. Estimator:             Unadjusted
                                        F-statistic:                     7.6332
Entities:                         128   P-value                          0.0000
Avg Obs:                       4.0000   Distribution:                  F(8,503)
Min Obs:                       4.0000
Max Obs:                       4.0000   F-statistic (robust):            7.6332
                                        P-value                          0.0000
Time periods:                       4   Distribution:                  F(8,503)
Avg Obs:                      128.00
Min Obs:                      128.00
Max Obs:                      128.00

                                        Parameter Estimates
==============================================================================
====================
                                 Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
------------------------------------------------------------------------------
--------------------
const                               13.455     0.2902     46.365     0.0000
12.885      14.025
Asset Turnover                     -0.5035     0.1431    -3.5180     0.0005
-0.7847     -0.2223
Current Ratio                       0.0391     0.1079     0.3625     0.7171
-0.1728      0.2511
Enterprise Value to EBIT           -0.0011     0.0010    -1.1116     0.2668
```

```
-0.0031      0.0009
Interest Coverage Ratio             -0.0005     0.0002    -2.2680     0.0238
-0.0010  -7.027e-05
Price to Cash Flow per Share        -0.0383     0.0086    -4.4648     0.0000
-0.0552     -0.0215
Return on Assets                    -0.0194     0.0155    -1.2557     0.2098
-0.0498      0.0110
Total Debt to Total Equity           0.0006     0.0006     0.9985     0.3185
-0.0006      0.0018
Weighted Average Cost of Capital    -0.0168     0.0371    -0.4522     0.6513
-0.0897      0.0561
==============================================================================
====================
```

Model 3 – Scope 2 Market

[51]:
```python
# create Y df
Y3 = np.log(final_data[['Scope2 Market']])

# specify the model
mod3 = PanelOLS(Y3, X)
res3 = mod3.fit()
print(res3)
```

```
                          PanelOLS Estimation Summary
==============================================================================
Dep. Variable:          Scope2 Market   R-squared:                       0.1110
Estimator:                   PanelOLS   R-squared (Between):             0.1329
No. Observations:                 512   R-squared (Within):             -0.1193
Date:                Thu, Jun 15 2023   R-squared (Overall):             0.1110
Time:                        13:58:43   Log-likelihood                  -1105.4
Cov. Estimator:            Unadjusted
                                        F-statistic:                     7.8467
Entities:                         128   P-value                          0.0000
Avg Obs:                       4.0000   Distribution:                  F(8,503)
Min Obs:                       4.0000
Max Obs:                       4.0000   F-statistic (robust):            7.8467
                                        P-value                          0.0000
Time periods:                       4   Distribution:                  F(8,503)
Avg Obs:                      128.00
Min Obs:                      128.00
Max Obs:                      128.00

                                        Parameter Estimates
==============================================================================
====================
```

```
                                 Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
const                                12.270     0.3868     31.718     0.0000
11.510      13.030
Asset Turnover                      -0.0957     0.1908    -0.5013     0.6164
-0.4705      0.2792
Current Ratio                        0.0197     0.1438     0.1367     0.8914
-0.2629      0.3022
Enterprise Value to EBIT            -0.0012     0.0013    -0.8605     0.3899
-0.0038      0.0015
Interest Coverage Ratio             -0.0001     0.0003    -0.4232     0.6723
-0.0007      0.0005
Price to Cash Flow per Share        -0.0407     0.0114    -3.5590     0.0004
-0.0632     -0.0183
Return on Assets                    -0.0938     0.0206    -4.5477     0.0000
-0.1344     -0.0533
Total Debt to Total Equity           0.0011     0.0008     1.3771     0.1691
-0.0005      0.0026
Weighted Average Cost of Capital     0.0909     0.0495     1.8382     0.0666
-0.0063      0.1881
================================================================================
====================
```

Model 4 – Sum of Scope1 and Scope2 Location

[52]:
```python
# create Y df
Y4 = np.log(final_data[['Sum_Emission_Loc']])

# specify the model
mod4 = PanelOLS(Y4, X)
res4 = mod4.fit()
print(res4)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:       Sum_Emission_Loc   R-squared:                        0.1624
Estimator:                   PanelOLS   R-squared (Between):              0.1966
No. Observations:                 512   R-squared (Within):              -2.8869
Date:                Thu, Jun 15 2023   R-squared (Overall):              0.1624
Time:                        13:58:43   Log-likelihood                   -1009.9
Cov. Estimator:            Unadjusted
                                        F-statistic:                      12.186
Entities:                        128   P-value                           0.0000
Avg Obs:                      4.0000   Distribution:                   F(8,503)
```

```
Min Obs:                        4.0000
Max Obs:                        4.0000   F-statistic (robust):             12.186
                                         P-value                           0.0000
Time periods:                        4   Distribution:                   F(8,503)
Avg Obs:                        128.00
Min Obs:                        128.00
Max Obs:                        128.00

                                          Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
const                                14.871     0.3210     46.322     0.0000
14.240      15.502
Asset Turnover                      -0.6127     0.1583    -3.8696     0.0001
-0.9238     -0.3016
Current Ratio                        0.0019     0.1193     0.0162     0.9871
-0.2325      0.2364
Enterprise Value to EBIT            -0.0022     0.0011    -2.0118     0.0448
-0.0044  -5.257e-05
Interest Coverage Ratio             -0.0006     0.0003    -2.3110     0.0212
-0.0011  -8.873e-05
Price to Cash Flow per Share        -0.0458     0.0095    -4.8196     0.0000
-0.0645     -0.0271
Return on Assets                    -0.0438     0.0171    -2.5575     0.0108
-0.0774     -0.0102
Total Debt to Total Equity           0.0007     0.0007     1.0850     0.2784
-0.0006      0.0020
Weighted Average Cost of Capital    -0.0562     0.0410    -1.3683     0.1718
-0.1368      0.0245
================================================================================
====================
```

Model 5 – Sum of Scope1 and Scope2 Market

[53]:
```python
# create Y df
Y5 = np.log(final_data[['Sum_Emission_Mkt']])

# specify the model
mod5 = PanelOLS(Y5, X)
res5 = mod5.fit()
print(res5)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:       Sum_Emission_Mkt   R-squared:                        0.1649
Estimator:                   PanelOLS   R-squared (Between):              0.1996
No. Observations:                 512   R-squared (Within):              -1.5943
Date:                Thu, Jun 15 2023   R-squared (Overall):              0.1649
Time:                        13:58:43   Log-likelihood                   -1049.9
Cov. Estimator:            Unadjusted
                                        F-statistic:                      12.413
Entities:                         128   P-value                           0.0000
Avg Obs:                       4.0000   Distribution:                   F(8,503)
Min Obs:                       4.0000
Max Obs:                       4.0000   F-statistic (robust):             12.413
                                        P-value                           0.0000
Time periods:                       4   Distribution:                   F(8,503)
Avg Obs:                      128.00
Min Obs:                      128.00
Max Obs:                      128.00

                                        Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
const                                14.564     0.3471     41.964     0.0000
13.882      15.246
Asset Turnover                      -0.4544     0.1712    -2.6544     0.0082
-0.7907     -0.1181
Current Ratio                       -0.0147     0.1290    -0.1143     0.9090
-0.2682      0.2387
Enterprise Value to EBIT            -0.0027     0.0012    -2.2205     0.0268
-0.0050     -0.0003
Interest Coverage Ratio             -0.0005     0.0003    -1.6654     0.0965
-0.0010   8.291e-05
Price to Cash Flow per Share        -0.0480     0.0103    -4.6745     0.0000
-0.0682     -0.0278
Return on Assets                    -0.0704     0.0185    -3.8011     0.0002
-0.1067     -0.0340
Total Debt to Total Equity           0.0007     0.0007     1.0541     0.2924
-0.0006      0.0021
Weighted Average Cost of Capital    -0.0381     0.0444    -0.8582     0.3912
-0.1252      0.0491
================================================================================
====================
```

Model Comparison

[54]:
```
print(compare({"Model 1": res1,
               "Model 2": res2,
               "Model 3": res3,
               "Model 4": res4,
               "Model 5": res5}, stars=True).summary)
```

```
                                                    Model Comparison
================================================================================
=====================================
                                          Model 1           Model 2         Model
3            Model 4            Model 5
--------------------------------------------------------------------------------
-------------------------------------
Dep. Variable                              Scope1   Scope2 Location   Scope2
Market   Sum_Emission_Loc   Sum_Emission_Mkt
Estimator                                 PanelOLS          PanelOLS
PanelOLS           PanelOLS           PanelOLS
No. Observations                               512               512
512                512                512
Cov. Est.                               Unadjusted        Unadjusted
Unadjusted         Unadjusted         Unadjusted
R-squared                                   0.1787            0.1083
0.1110             0.1624             0.1649
R-Squared (Within)                         -4.2813           -1.2580
-0.1193            -2.8869            -1.5943
R-Squared (Between)                         0.2161            0.1348
0.1329             0.1966             0.1996
R-Squared (Overall)                         0.1787            0.1083
0.1110             0.1624             0.1649
F-statistic                                 13.678            7.6332
7.8467             12.186             12.413
P-value (F-stat)                            0.0000            0.0000
0.0000             0.0000             0.0000
================================== ============ =================
=============== ================== ==================
const                                   14.487***         13.455***
12.270***          14.871***          14.564***
                                         (36.230)          (46.365)
(31.718)           (46.322)           (41.964)
Asset Turnover                         -0.6201***        -0.5035***
-0.0957         -0.6127***         -0.4544***
                                        (-3.1440)         (-3.5180)
(-0.5013)          (-3.8696)          (-2.6544)
Current Ratio                           -0.1906            0.0391
0.0197             0.0019            -0.0147
                                        (-1.2819)          (0.3625)
```

```
(0.1367)           (0.0162)          (-0.1143)
Enterprise Value to EBIT             -0.0048***           -0.0011
-0.0012          -0.0022**          -0.0027**
                                      (-3.4329)         (-1.1116)
(-0.8605)          (-2.0118)          (-2.2205)
Interest Coverage Ratio               -0.0005*         -0.0005**
-0.0001          -0.0006**          -0.0005*
                                      (-1.7226)         (-2.2680)
(-0.4232)          (-2.3110)          (-1.6654)
Price to Cash Flow per Share         -0.0561***        -0.0383***
-0.0407***         -0.0458***         -0.0480***
                                      (-4.7436)         (-4.4648)
(-3.5590)          (-4.8196)          (-4.6745)
Return on Assets                     -0.0648***           -0.0194
-0.0938***          -0.0438**         -0.0704***
                                      (-3.0393)         (-1.2557)
(-4.5477)          (-2.5575)          (-3.8011)
Total Debt to Total Equity              0.0001            0.0006
0.0011             0.0007             0.0007
                                       (0.1323)          (0.9985)
(1.3771)           (1.0850)           (1.0541)
Weighted Average Cost of Capital       -0.0717           -0.0168
0.0909*            -0.0562            -0.0381
                                      (-1.4017)         (-0.4522)
(1.8382)          (-1.3683)          (-0.8582)
--------------------------------------------------------------------------------
---------------------------------------

T-stats reported in parentheses
```

[55]:
```
summary1 = compare({"Model 1": res1,
                "Model 2": res2,
                "Model 3": res3,
                "Model 4": res4,
                "Model 5": res5}, stars=True).summary
with open('Regression_Summary1.csv','w') as f:
    f.write(summary1.as_csv())
```

2.2.2 METHOD 2

[56]:
```
final_data.columns =␣
↪['Sector','Scope1','Scope2_Location','Scope2_Market','Sum_Emission_Mkt','Sum_Emission_Loc',
                     'Asset_Turnover', 'Current_Ratio',␣
↪'Enterprise_Value_to_EBIT', 'Interest_Coverage_Ratio',
                     'Price_to_Cash_Flow_per_Share', 'Return_on_Assets' ,␣
↪'Total_Debt_to_Total_Equity',
                     'Weighted_Average_Cost_of_Capital']
```

Model 1 – Scope 1

[57]:
```
model1 = PanelOLS.from_formula('log(Scope1) ~ Asset_Turnover + Current_Ratio +␣
↪Enterprise_Value_to_EBIT + Interest_Coverage_Ratio +␣
↪Price_to_Cash_Flow_per_Share + Return_on_Assets + Total_Debt_to_Total_Equity␣
↪+ Weighted_Average_Cost_of_Capital', data = final_data)
results1 = model1.fit(cov_type='clustered', cluster_entity=True)
print(results1)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:            log(Scope1)   R-squared:                       0.8843
Estimator:                   PanelOLS   R-squared (Between):             0.8989
No. Observations:                 512   R-squared (Within):             -43.927
Date:                Thu, Jun 15 2023   R-squared (Overall):             0.8843
Time:                        13:58:43   Log-likelihood                  -1451.0
Cov. Estimator:             Clustered
                                        F-statistic:                     481.71
Entities:                         128   P-value                          0.0000
Avg Obs:                       4.0000   Distribution:                  F(8,504)
Min Obs:                       4.0000
Max Obs:                       4.0000   F-statistic (robust):            232.02
                                        P-value                          0.0000
Time periods:                       4   Distribution:                  F(8,504)
Avg Obs:                      128.00
Min Obs:                      128.00
Max Obs:                      128.00

                                        Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
Asset_Turnover                       2.0125     0.7082     2.8417     0.0047
0.6211      3.4039
Current_Ratio                        1.9672     0.4322     4.5516     0.0000
1.1181      2.8164
Enterprise_Value_to_EBIT             0.0007     0.0017     0.4154     0.6780
-0.0026      0.0040
Interest_Coverage_Ratio             -0.0006     0.0004    -1.6387     0.1019
-0.0013      0.0001
Price_to_Cash_Flow_per_Share        -0.0001     0.0270    -0.0045     0.9964
-0.0531      0.0528
Return_on_Assets                    -0.0047     0.0511    -0.0920     0.9267
-0.1051      0.0957
Total_Debt_to_Total_Equity           0.0127     0.0066     1.9216     0.0552
```

```
-0.0003      0.0256
Weighted_Average_Cost_of_Capital     0.9177     0.1030     8.9104     0.0000
0.7154      1.1201
================================================================================
====================
```

Model 2 – Scope2 Location

[58]:
```
model2 = PanelOLS.from_formula('log(Scope2_Location) ~ Asset_Turnover +␣
↪Current_Ratio + Enterprise_Value_to_EBIT + Interest_Coverage_Ratio +␣
↪Price_to_Cash_Flow_per_Share + Return_on_Assets + Total_Debt_to_Total_Equity␣
↪+ Weighted_Average_Cost_of_Capital', data = final_data)
results2 = model2.fit(cov_type='clustered', cluster_entity=True)
print(results2)
```

```
                          PanelOLS Estimation Summary
================================================================================
==
Dep. Variable:     log(Scope2_Location)   R-squared:
0.9156
Estimator:                     PanelOLS   R-squared (Between):
0.9292
No. Observations:                   512   R-squared (Within):
-38.764
Date:                  Thu, Jun 15 2023   R-squared (Overall):
0.9156
Time:                          13:58:43   Log-likelihood
-1383.9
Cov. Estimator:               Clustered
                                          F-statistic:
683.55
Entities:                           128   P-value
0.0000
Avg Obs:                         4.0000   Distribution:
F(8,504)
Min Obs:                         4.0000
Max Obs:                         4.0000   F-statistic (robust):
386.81
                                          P-value
0.0000
Time periods:                         4   Distribution:
F(8,504)
Avg Obs:                        128.00
Min Obs:                        128.00
Max Obs:                        128.00
```

```
                                        Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
Asset_Turnover                       1.9415     0.6505     2.9848     0.0030
0.6635      3.2194
Current_Ratio                        2.0431     0.3273     6.2419     0.0000
1.4000      2.6862
Enterprise_Value_to_EBIT             0.0040     0.0016     2.4095     0.0163
0.0007      0.0072
Interest_Coverage_Ratio             -0.0006     0.0003    -2.0266     0.0432
-0.0011  -1.725e-05
Price_to_Cash_Flow_per_Share         0.0137     0.0227     0.6038     0.5463
-0.0308      0.0582
Return_on_Assets                     0.0364     0.0463     0.7857     0.4324
-0.0546      0.1274
Total_Debt_to_Total_Equity           0.0122     0.0063     1.9349     0.0536
-0.0002      0.0247
Weighted_Average_Cost_of_Capital     0.9021     0.0969     9.3089     0.0000
0.7117      1.0925
================================================================================
====================
```

Model 3 – Scope2 Market

[59]:
```
model3 = PanelOLS.from_formula('log(Scope2_Market) ~ Asset_Turnover +␣
↪Current_Ratio + Enterprise_Value_to_EBIT + Interest_Coverage_Ratio +␣
↪Price_to_Cash_Flow_per_Share + Return_on_Assets + Total_Debt_to_Total_Equity␣
↪+ Weighted_Average_Cost_of_Capital', data = final_data)
results3 = model3.fit(cov_type='clustered', cluster_entity=True)
print(results3)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:     log(Scope2_Market)   R-squared:                        0.9047
Estimator:                   PanelOLS   R-squared (Between):              0.9197
No. Observations:                 512   R-squared (Within):              -3.9114
Date:                Thu, Jun 15 2023   R-squared (Overall):              0.9047
Time:                        13:58:43   Log-likelihood                   -1386.7
Cov. Estimator:             Clustered
                                        F-statistic:                      598.26
Entities:                         128   P-value                           0.0000
Avg Obs:                       4.0000   Distribution:                   F(8,504)
```

```
Min Obs:                        4.0000
Max Obs:                        4.0000   F-statistic (robust):             423.76
                                         P-value                           0.0000
Time periods:                        4   Distribution:                   F(8,504)
Avg Obs:                       128.00
Min Obs:                       128.00
Max Obs:                       128.00

                                         Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
Asset_Turnover                       2.1341     0.6598     3.2344     0.0013
0.8378      3.4304
Current_Ratio                        1.8472     0.3389     5.4509     0.0000
1.1814      2.5130
Enterprise_Value_to_EBIT             0.0035     0.0019     1.8376     0.0667
-0.0002      0.0072
Interest_Coverage_Ratio             -0.0002     0.0002    -0.6682     0.5043
-0.0007      0.0003
Price_to_Cash_Flow_per_Share         0.0067     0.0239     0.2802     0.7794
-0.0403      0.0537
Return_on_Assets                    -0.0429     0.0570    -0.7529     0.4519
-0.1549      0.0691
Total_Debt_to_Total_Equity           0.0117     0.0059     1.9851     0.0477
0.0001      0.0233
Weighted_Average_Cost_of_Capital     0.9289     0.0972     9.5588     0.0000
0.7380      1.1198
================================================================================
====================
```

Model 4 – Sum of Scope 1 and Scope2 Location

[60]:
```
model4 = PanelOLS.from_formula('log(Sum_Emission_Loc) ~ Asset_Turnover +␣
↪Current_Ratio + Enterprise_Value_to_EBIT + Interest_Coverage_Ratio +␣
↪Price_to_Cash_Flow_per_Share + Return_on_Assets + Total_Debt_to_Total_Equity␣
↪+ Weighted_Average_Cost_of_Capital', data = final_data)
results4 = model4.fit(cov_type='clustered', cluster_entity=True)
print(results4)
```

```
                          PanelOLS Estimation Summary
================================================================================
===
```

```
Dep. Variable:     log(Sum_Emission_Loc)   R-squared:
0.9084
Estimator:                      PanelOLS   R-squared (Between):
0.9224
No. Observations:                    512   R-squared (Within):
-59.504
Date:                   Thu, Jun 15 2023   R-squared (Overall):
0.9084
Time:                           13:58:43   Log-likelihood
-1435.2
Cov. Estimator:                Clustered
                                           F-statistic:
625.03
Entities:                            128   P-value
0.0000
Avg Obs:                          4.0000   Distribution:
F(8,504)
Min Obs:                          4.0000
Max Obs:                          4.0000   F-statistic (robust):
357.23
                                           P-value
0.0000
Time periods:                          4   Distribution:
F(8,504)
Avg Obs:                          128.00
Min Obs:                          128.00
Max Obs:                          128.00

                                       Parameter Estimates
================================================================================
====================
                                  Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
Asset_Turnover                       2.0896     0.6983     2.9924     0.0029
0.7177      3.4616
Current_Ratio                        2.2169     0.3819     5.8043     0.0000
1.4665      2.9673
Enterprise_Value_to_EBIT             0.0034     0.0017     1.9802     0.0482
2.631e-05      0.0067
Interest_Coverage_Ratio             -0.0006     0.0003    -2.0110     0.0449
-0.0013  -1.466e-05
Price_to_Cash_Flow_per_Share         0.0117     0.0248     0.4717     0.6374
-0.0371      0.0605
Return_on_Assets                     0.0179     0.0487     0.3682     0.7129
-0.0777      0.1135
Total_Debt_to_Total_Equity           0.0136     0.0070     1.9389     0.0531
```

```
-0.0002      0.0274
Weighted_Average_Cost_of_Capital     0.9594     0.1018     9.4285     0.0000
0.7595      1.1594
==============================================================================
====================
```

Model 5 – Sum of Scope 1 and Scope2 Market

[61]:
```python
model5 = PanelOLS.from_formula('log(Sum_Emission_Mkt) ~ Asset_Turnover +␣
↪Current_Ratio + Enterprise_Value_to_EBIT + Interest_Coverage_Ratio +␣
↪Price_to_Cash_Flow_per_Share + Return_on_Assets + Total_Debt_to_Total_Equity␣
↪+ Weighted_Average_Cost_of_Capital', data = final_data)
results5 = model5.fit(cov_type='clustered', cluster_entity=True)
print(results5)
```

```
                          PanelOLS Estimation Summary
==============================================================================
===
Dep. Variable:     log(Sum_Emission_Mkt)   R-squared:
0.9042
Estimator:                       PanelOLS   R-squared (Between):
0.9181
No. Observations:                     512   R-squared (Within):
-27.240
Date:                   Thu, Jun 15 2023   R-squared (Overall):
0.9042
Time:                           13:58:44   Log-likelihood
-1435.0
Cov. Estimator:                Clustered
                                           F-statistic:
594.79
Entities:                             128   P-value
0.0000
Avg Obs:                           4.0000   Distribution:
F(8,504)
Min Obs:                           4.0000
Max Obs:                           4.0000   F-statistic (robust):
312.91
                                           P-value
0.0000
Time periods:                           4   Distribution:
F(8,504)
Avg Obs:                          128.00
Min Obs:                          128.00
Max Obs:                          128.00
```

```
                                        Parameter Estimates
================================================================================
====================
                                 Parameter  Std. Err.     T-stat    P-value
Lower CI    Upper CI
--------------------------------------------------------------------------------
--------------------
Asset_Turnover                      2.1922     0.7086     3.0938     0.0021
0.8001      3.5843
Current_Ratio                       2.1545     0.3942     5.4649     0.0000
1.3799      2.9291
Enterprise_Value_to_EBIT            0.0028     0.0018     1.5892     0.1126
-0.0007      0.0063
Interest_Coverage_Ratio            -0.0005     0.0003    -1.6231     0.1052
-0.0011      0.0001
Price_to_Cash_Flow_per_Share        0.0083     0.0250     0.3326     0.7395
-0.0407      0.0573
Return_on_Assets                   -0.0099     0.0509    -0.1949     0.8455
-0.1100      0.0901
Total_Debt_to_Total_Equity          0.0134     0.0068     1.9599     0.0506
-3.248e-05      0.0268
Weighted_Average_Cost_of_Capital    0.9566     0.1013     9.4408     0.0000
0.7575      1.1556
================================================================================
====================
```

2.2.3 Model Comparison

[62]:
```
compare({"Model 1": results1,
         "Model 2": results2,
         "Model 3": results3,
         "Model 4": results4,
         "Model 5": results5}, stars=True).summary
```

[62]:
```
<class 'linearmodels.compat.statsmodels.Summary'>
"""
                                                                Model Comparison
================================================================================
==============================================================
                                        Model 1                Model 2
Model 3                 Model 4                 Model 5
--------------------------------------------------------------------------------
--------------------------------------------------------------
Dep. Variable                       log(Scope1)   log(Scope2_Location)
log(Scope2_Market)   log(Sum_Emission_Loc)   log(Sum_Emission_Mkt)
```

Estimator	PanelOLS	PanelOLS	PanelOLS	PanelOLS	PanelOLS
No. Observations	512	512	512	512	512
Cov. Est.	Clustered	Clustered	Clustered	Clustered	Clustered
R-squared	0.8843	0.9156	0.9047	0.9084	0.9042
R-Squared (Within)	-43.927	-38.764	-3.9114	-59.504	-27.240
R-Squared (Between)	0.8989	0.9292	0.9197	0.9224	0.9181
R-Squared (Overall)	0.8843	0.9156	0.9047	0.9084	0.9042
F-statistic	481.71	683.55	598.26	625.03	594.79
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000	0.0000
Asset_Turnover	2.0125***	1.9415***	2.1341***	2.0896***	2.1922***
	(2.8417)	(2.9848)	(3.2344)	(2.9924)	(3.0938)
Current_Ratio	1.9672***	2.0431***	1.8472***	2.2169***	2.1545***
	(4.5516)	(6.2419)	(5.4509)	(5.8043)	(5.4649)
Enterprise_Value_to_EBIT	0.0007	0.0040**	0.0035*	0.0034**	0.0028
	(0.4154)	(2.4095)	(1.8376)	(1.9802)	(1.5892)
Interest_Coverage_Ratio	-0.0006	-0.0006**	-0.0002	-0.0006**	-0.0005
	(-1.6387)	(-2.0266)	(-0.6682)	(-2.0110)	(-1.6231)
Price_to_Cash_Flow_per_Share	-0.0001	0.0137	0.0067	0.0117	0.0083
	(-0.0045)	(0.6038)	(0.2802)	(0.4717)	(0.3326)
Return_on_Assets	-0.0047	0.0364	-0.0429	0.0179	-0.0099
	(-0.0920)	(0.7857)	(-0.7529)	(0.3682)	(-0.1949)
Total_Debt_to_Total_Equity	0.0127*	0.0122*	0.0117**	0.0136*	0.0134*
	(1.9216)	(1.9349)			

```
(1.9851)                (1.9389)                (1.9599)
Weighted_Average_Cost_of_Capital       0.9177***               0.9021***
0.9289***               0.9594***               0.9566***
                                         (8.9104)                (9.3089)
(9.5588)                (9.4285)                (9.4408)
--------------------------------------------------------------------------------
-------------------------------------------------------------

T-stats reported in parentheses
"""
```

[63]:
```
summary2 = compare({"Model 1": results1,
                    "Model 2": results2,
                    "Model 3": results3,
                    "Model 4": results4,
                    "Model 5": results5}, stars=True).summary
with open('Regression_Summary2.csv','w') as f:
    f.write(summary2.as_csv())
```