Project-1 Tina

2024-10-20

Abstract

The dataset for this project is about the ESG (Environmental, Social, and Governance) scores of S&P 500 companies in 2023-2024. The project attempts to find the relationship between different industries and their ESG scores. The goal is to further explore the industries that focus on ESG index, compare the distribution of information technology, healthcare, and finance in ESG index, and explore the importance that different industries attach to governance scores. As a result, Real Estate, Information Technology, and Consumer Discretionary are the top three industries that focus on ESG index. The Information Technology industry is more conscious of its impact than the Financial and Health Care industries. Surprisingly, the Financial sector has a relatively higher governance score than all other sectors.

Introduction

As society develops, the world is paying more and more attention to the performance and impact of companies in terms of environment, society, and governance. As leaders and models in various industries, the S&P 500 companies impact other companies worldwide. Their performance and management direction play a leading role. Different sectors may pay distinct attention to the ESG index. This project mainly explores the ESG scores of various industries and their emphasis on ESG impact.

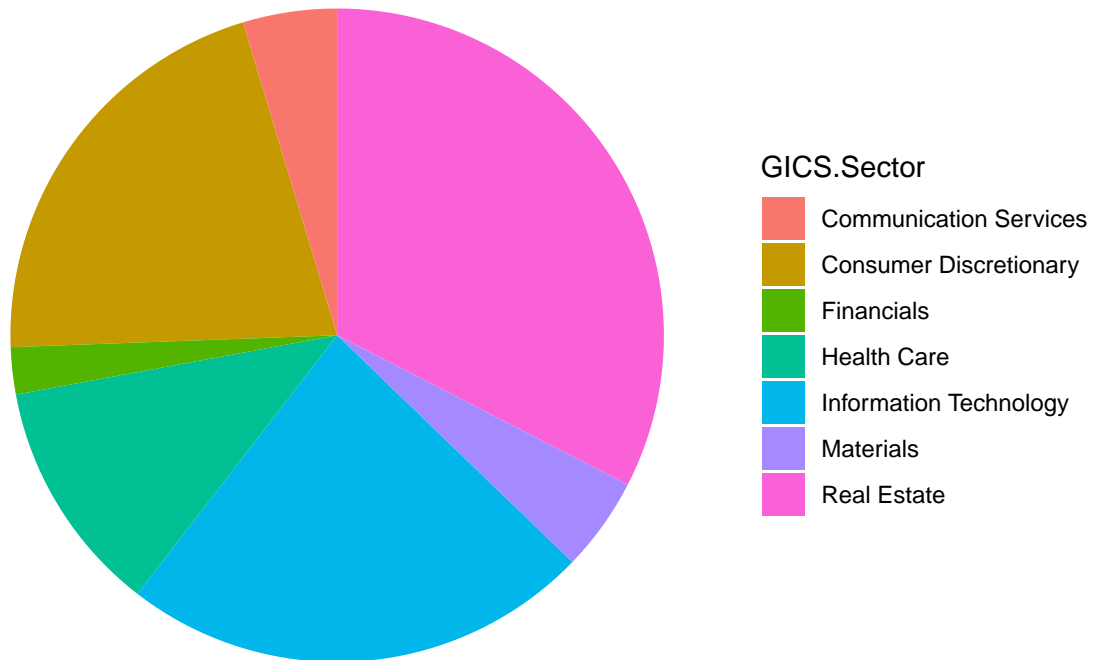## Industry Distribution of the 10% Companies with the Lowest ESG Scores

This section aims to investigate the 10% of companies with the lowest ESG total scores and their industry distribution to explore the correlation between score and type of industry.

```
# Identify top 10% companies with lowest score
top_10_percent_value <- quantile(data$totalEsg, probs=0.10)
companies_with_lowest_score <- data |> filter(totalEsg <= top_10_percent_value)

# Create a pie chart to explore the distribution of industries in the 90th percentile
sector_counts <- companies_with_lowest_score |>
  group_by(GICS.Sector) |>
  summarise(count = n()) |>
  arrange(desc(count))

ggplot(sector_counts, aes(x = "", y = count, fill = GICS.Sector)) +
  geom_bar(stat = "identity") +
  coord_polar("y", start = 0) +
  theme_void() +
  labs(title = "Industry Distribution in the 90th Percentile of ESG Scores")
```

Industry Distribution in the 90th Percentile of ESG Scores


```
print(sector_counts)
```

```
## # A tibble: 7 x 2
##   GICS.Sector              count
##   <chr>                   <int>
## 1 Real Estate                14
## 2 Information Technology     10
## 3 Consumer Discretionary      9
## 4 Health Care                 5
## 5 Communication Services      2
## 6 Materials                   2
## 7 Financials                  1
```

The pie chart shows that among the lowest 10% of the S&P 500 ESG index, the real estate industry has the most companies, 14. The information technology and Consumer Discretionary industries are second and third. This reflects that society has attached great importance to the real estate industry's concept of green environmental protection in recent years.