

# ESG Beyond Words: Leveraging Text Classification to Measure Corporate Commitment

*Carol Schuetz, Business Librarian, Baylor University*

*Millicent Weber, Data Science Librarian, Baylor University*

*Joshua Been, Director of Data and Digital Scholarship, Baylor University*

## What We Did

- Text mining Annual Reports and 10-K reports
- What are companies saying about ESG topics?

## Problem We are Trying to Solve

- ▶ How to measure ESG engagement in a timely fashion
- ▶ How to measure the level of ESG engagement

# ESG Buckets

<b>Environmental</b>	<b>Social</b>	<b>Governance</b>
Air Quality	Employee diversity	Board Member Diversity
Water Pollution	Data security	Political Contributions
Deforestation	Customer satisfaction	Executive Pay
Green Energy Initiatives	Sexual harassment policies	Lawsuits
Waste Management	Human trafficking and slavery	Internal Corruption
Water Usage	Working conditions	Lobbying
Resource Depletion	Global community service	Tax Strategy
Climate Change	Local community service	Venture Partner Compensation
Environmental Sustainability	Employee Health	Hiring Policies
		Onboarding Policies

5

## Identifying ESG-Related Sentences in Corporate Annual Reports



Classification Methods



Publicly Available Jupyter Notebook on Google Colab



Flow of the Python Code on Colab



Output



Accuracy Rate

# How can a computer identify this sentence as *Environmental*?

As planetary health impacts public health, we are committed to limiting our impact on the environment and the climate. Our company purpose – breakthroughs that change patients' lives – guides our environmental priorities, with a focus on impact reduction, conservation of resources and the reduction of waste arising from our operations. As a global biopharmaceutical company, we are committed to applying science to fight the challenges posed by the climate crisis and innovating to bring products to the market that tackle health conditions linked to our warming planet.



Pfizer Annual Report (2021)

# Classification Methods

## Discarded Methods

- Topic Modeling

*Terms are grouped into topics with no regard to context*

- Latent Dirichlet Allocation (LDA)
- Latent Semantic Analysis (LSA)

- Principal Component Analysis (PCA)

*Takes a set of possibly correlated variables and transforms them into uncorrelated variables called principal components*

*Risk of data loss due to reducing the dimensionality of the content.*

# Popular Classification Methods

## We Implemented: Zero-Shot Classification

- Artificial Intelligence
  - Based on Large Language Models (LLM)
  - Train an AI on a large quantity of text
- Sentence Classification
  - Provide the LLM with a new sentence and it will compare it to its understanding of the text it was trained on and it will classify it.
- Classification Descriptions
  - Descriptions for each category can be provided to help the LLM to classify sentences.





# (JB) Zero-Shot Classification

# Public LLM on Hugging Face: bart-large-mnli

BART

- Bidirectional and Auto-Regressive Transformers
- Method used to train this model

Large

- This is a Large Language Model (LLM)

Multi-Genre Natural Language Inference (MultiNLI)

- Crowd-sourced collection of 433k sentence pairs annotated with textual entailment information.

The screenshot shows the Hugging Face model card for 'facebook/bart-large-mnli'. The page includes a search bar, navigation tabs for Models, Datasets, Spaces, Docs, and Solutions, and a search bar for models, datasets, and users. The model card itself features a title 'facebook/bart-large-mnli' with 713 likes, a license of 'mit', and a 'Model card' tab. The main content area is titled 'bart-large-mnli' and describes it as the checkpoint for 'bart-large' trained on the 'MultiNLI (MNLI)' dataset. It provides additional information with links to the 'bart-large' model page, the paper 'BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension', and the 'BART fairseq implementation'. A section titled 'NLI-based Zero Shot Text Classification' explains the method proposed by Yin et al., which uses pre-trained NLI models as zero-shot sequence classifiers. The text describes how a hypothesis is constructed from candidate labels and how probabilities for entailment and contradiction are converted to label probabilities. It also mentions the effectiveness of this method with larger pre-trained models like BART and Roberta, and provides a link to a blog post for more details. On the right side, there is a 'Hosted inference API' section with a 'Zero-Shot Classification' example. The example shows a text input: 'Last week I upgraded my iOS version and ever since my phone has been overheating whenever I use your app.' The possible class names are 'mobile, website, billing, account access'. The 'Compute' button is highlighted, and the output shows the predicted class: 'mobile'.

Google Colab  
is a free  
service to  
execute and  
share Python  
and R scripts

```
62
63 try:
64     from transformers import pipeline
65 except:
66     !pip install tokenizers &> /dev/null
67     !git clone https://github.com/huggingface/transformers &> /dev/null
68     !pip install transformers/. &> /dev/null
69     from transformers import pipeline
70
71 pbar.update(1)
72
73 classifier = pipeline("zero-shot-classification", model='facebook/bart-large-mnli')
```

*Excerpt from script loading the LLM for classification*

# Google Colab: Code Block 1 Prepares Annual Reports

The screenshot displays a Google Colab notebook interface. The browser address bar shows the URL: `https://colab.research.google.com/drive/1qq4ouoFHIE3ze0-oQL8zKTrS9_0JNkbu#scrollTo=NN_N9aV7BsxS`. The notebook title is "Measure Commitment to Corporate ESG: Extracting from Annual Reports". The notebook content includes a code block that has executed, resulting in the following output:

**Extracting ESG from Annual Reports**

**Baylor University | Digital Scholarship**  
UNIVERSITY LIBRARIES

**SOURCE:** Google Drive

**Environmental\_Categories:** "air quality, water pollution, deforestation, green energy initiatives, waste management, water usage, resource depletion, climate change, environmental sustainability"

**Social\_Categories:** "employee diversity, data security, customer satisfaction, sexual harassment policies, human trafficking and slavery, working conditions, global community service, local community service, employee health"

**Governance\_Categories:** "board member diversity, political contributions, executive pay, lawsuits, internal corruption, lobbying, tax strategy, venture partner compensation, hiring policies, onboarding process"

**Max\_Sentences:** 100000

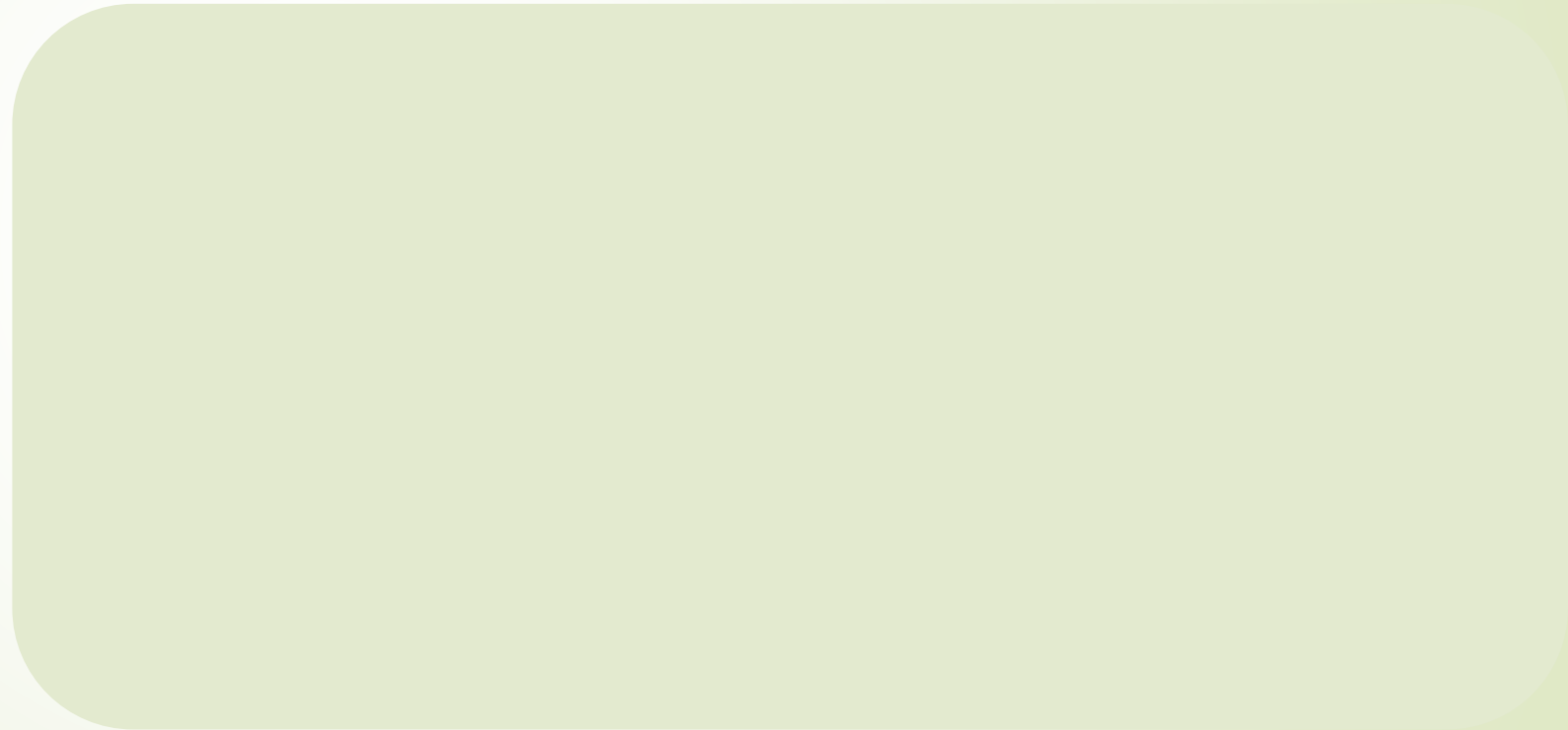
**Min\_Threshold:** 0.5

The output is displayed in a code block with a yellow circle highlighting the word "political" in the Governance\_Categories string. The status bar at the bottom indicates the execution completed in 3m 1s at 12:26 PM.

# Google Colab: Code Block 2 Flow

## Pre-Classification Test

- Is there a sufficient probability (0.5+) that a given sentence has a reference to a granular ESG bucket?
- Saves time so that we do not need to test every sentence for the 28 granular categories.



# Google Colab: Code Block 2 Flow

## Pre-Classification Test

- Is there a sufficient probability (0.5+) that a given sentence has a reference to a granular ESG bucket?
- Saves time so that we do not need to test every sentence for the 28 granular categories.



## Classification

- For sentences that pass the pre-test
- Sentences that match a granular ESG classification with a 0.5+ are recorded.



# Google Colab: Code Block 2 Flow

## Pre-Classification Test

- Is there a sufficient probability (0.5+) that a given sentence has a reference to a granular ESG bucket?
- Saves time so that we do not need to test every sentence for the 28 granular categories.



## Classification

- For sentences that pass the pre-test
- Sentences that match a granular ESG classification with a 0.5+ are recorded.



## Output

- Spreadsheet with each sentence and any classifications

# Code Block 2 Interface



SOURCE: Google Drive

[Show code](#)

Converting selected documents to text and saving in Google Drive. Next time, this step will be skipped.

100%  1/1 [00:00<00:00, 46.49it/s]

Cleaning texts using clean-text, regex, and spacy libraries. Each cleaned sentence is then passed into a list of sentences.

100%  1/1 [00:06<00:00, 6.47s/it]

**Environmental\_Categories:** " air quality, water pollution, deforestation, green energy initiatives, waste management, water usage, resource depletion, climate change, environmental sustainability "

**Social\_Categories:** " employee diversity, data security, customer satisfaction, sexual harassment policies, human trafficking and slavery, working conditions, global community service, local community service, employee health "

**Governance\_Categories:** " board member diversity, political contributions, executive pay, lawsuits, internal corruption, lobbying, tax strategy, venture partner compensation, hiring policies, onboarding process "

**Max\_Sentences:** 100000

**Min\_Threshold:** 0.5

[Show code](#)



A	B	C	K	L	
sentence	document	industry	climate change	environmental sustainability	employee
During 2021, we built on our previously announced plan to invest \$1.0 billion through the end of 2030 toward airline carbon neutrality (the "\$1.0 billion airline carbon neutrality goal") by announcing our intention to set new climate goals that are aligned with the applicable framework of the Science Based Targets initiative ("SBTi"): * Mediumterm goal: Reducing emissions intensity by 2035 as compared to 2019, based on jet fuel emissions of greenhouse gases ("GHGs") in accordance with SBTi guidance for the aviation industry, and in line with the Paris Agreement's goal of limiting global warming to well below two degrees Celsius above preindustrial levels.	Delta2021.pdf.txt	Airlines	0.640446842		-1
Longterm goal: Achieving net zero GHG emissions across the airline operation and its value chain (Scopes 1, 2 and 3) no later than 2050, as outlined by the SBTi Net Zero Standard Criteria.	Delta2021.pdf.txt	Airlines	-1	0.538130701	
We have submitted both goals to SBTi for validation, but we are unable to predict the outcome of that process and when it will be completed.	Delta2021.pdf.txt	Airlines	-1		-1
The global aviation industry is viewed as a hardtoabate sector, meaning it is innately difficult to decarbonize.	Delta2021.pdf.txt	Airlines	-1		-1
We expect our path toward achievement of these ambitious climate goals to include the following levers, with the fleet and fuel levers expected to play the most significant roles: Fleet: Currently, fleet					

# Output

# Accuracy Rate

We hand-coded (*classified*) sentences in 3 annual reports: Exxon, Ford, and Delta.

- ▶ True Positives (TP): Correctly identified by our model as ESG-related
  - ▶ 657
- ▶ True Negatives (TN): Correctly identified by our model as **not** ESG-related
  - ▶ 157,923
- ▶ False Positives (FP): **Incorrectly** identified by our model as ESG-related
  - ▶ 26
- ▶ False Negatives (FN): **Incorrectly** identified by our model as **not** ESG-related
  - ▶ 506

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

90+%

# Statistical Analysis Overview

- Scoring Metrics
  - Total Count
  - Average Score
  - Highest Score
  - 28-point Scoring System
- Comparing company to company, industry to industry

# Examples of Environmental Matches

E Bucket	Sentence	Score
<b>Air Quality</b>	When the U.S. government was deliberating on emissions control standards, Honda lobbied to set even higher standards on the basis of its superior technologies. (Honda 2021)	0.600
<b>Water Pollution</b>	Wiring harness clips in Ford Bronco Sport models are made of ocean-harvested plastic from discarded fishing nets. (Ford 2022)	0.506
<b>Green Energy Initiatives</b>	These projects, conducted at 15 of our hotels, are expected to generate average cost savings of 34% annually, while larger projects, such as chiller retrofits or replacements, are expected to generate energy cost savings of approximately 20% annually. (Park 2021)	0.640
<b>Waste Management</b>	Unfortunately, we saw another drop in the number of stores participating in recycling programs from 67% to 65%. (Roadhouse 2021)	0.699
<b>Resource Depletion</b>	As we produce crude oil and natural gas from our existing portfolio, the amount of our remaining reserves declines. (Conoco 2021)	0.723
<b>Climate Change</b>	We may be adversely affected by global climate change or by legal, regulatory or market responses to such change. (McKesson 2021)	0.873
<b>Environmental Sustainability</b>	JetBlue's primary environmental sustainability priority is reducing and managing carbon emissions from jet fuel. (JetBlue 2021)	0.943

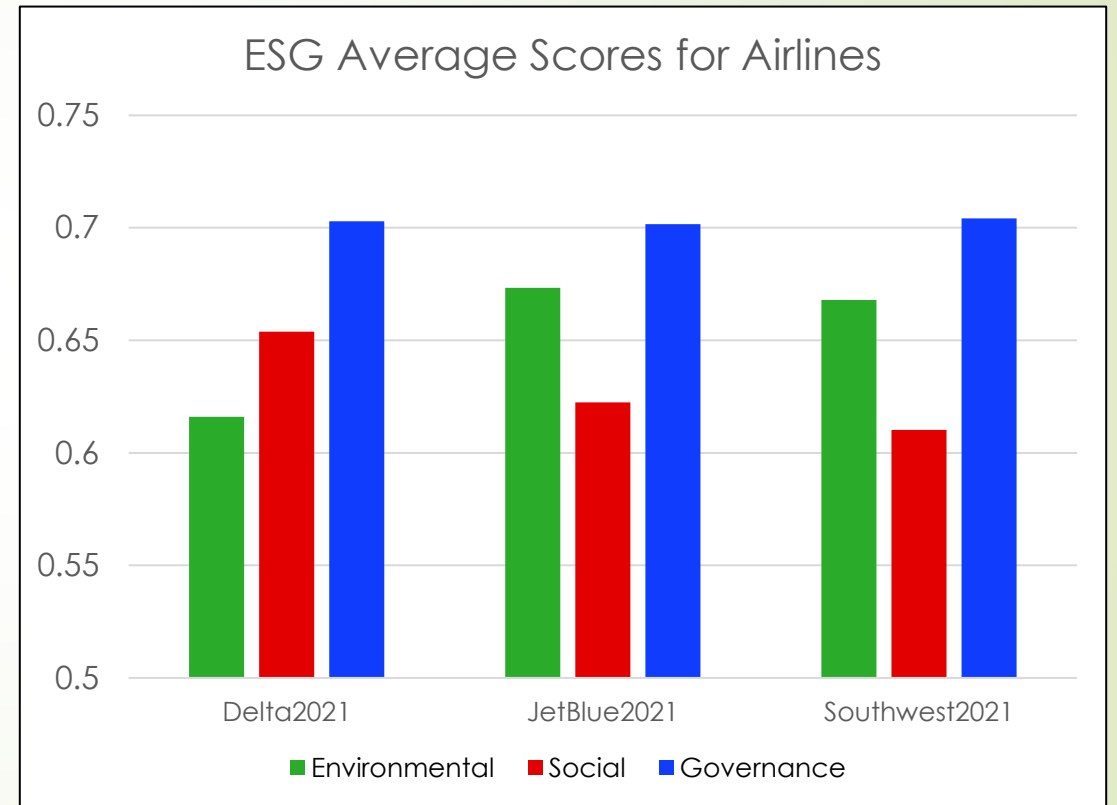
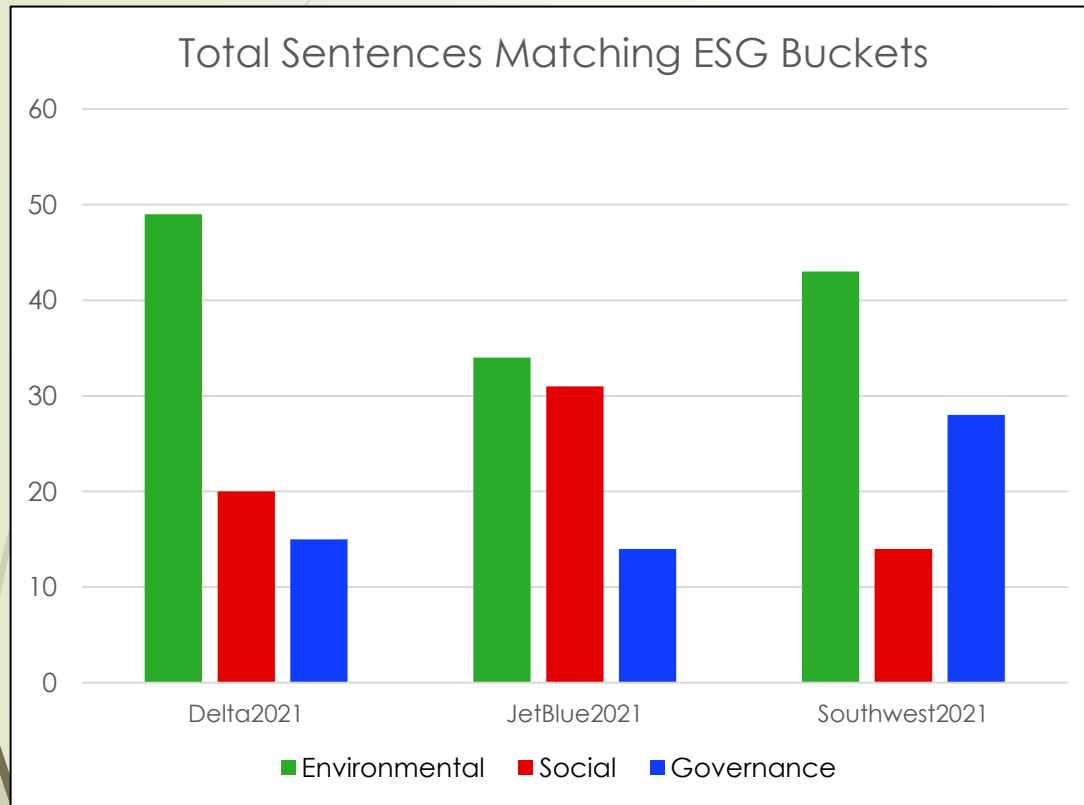
# Examples of Social Matches

S Bucket	Sentence	Score
<b>Employee diversity</b>	Our Employee Journeys research has yielded important insights about the experiences of women, Black/African American, Latinx, Asian, and LGBTQ employees at Lilly. (Eli Lilly 2020)	0.904
<b>Customer satisfaction</b>	In 2021, corporate travel professionals also rated us No. 1 in the annual Business Travel News Airline Survey for the 11th year in a row, citing our responsiveness to customers as well as our flexibility-minded approach throughout the pandemic and beyond.(Delta 2021)	0.652
<b>Human trafficking and slavery</b>	In 2021, we added a separate, mandatory training on diversity, equity, inclusion and unconscious bias for corporate employees, and we expanded our training on social issues to include modern slavery/human trafficking awareness. (Park 2021)	0.819
<b>Working conditions</b>	In addition, we have allowed our Support Center employees to maintain a work schedule that allows for working remotely on a periodic or fulltime basis depending on the prevalence of the virus. (Roadhouse 2021)	0.567
<b>Global community service</b>	We continued to focus on bringing water to communities around the world with our partnership with WaterStep, whose mission is to save lives with safe water. (Roadhouse 2021)	0.959
<b>Local community service</b>	Promoting progress in our hometown of Detroit through a \$50 million commitment to support nonprofits, education, employment and neighborhood revitalization efforts. (GM 2021)	0.955
<b>Employee Health</b>	In November 2021, Ford became the first U.S. automaker to require vaccinations for salaried employees. (Ford 2022)	0.654

# Examples of Governance Matches

G Bucket	Sentence	Score
<b>Board Member Diversity</b>	Additionally, the Company's Board of Directors has committed to increasing its diverse representation by 2025. (Southwest 2021)	0.870
<b>Political Contributions</b>	The board also exercises oversight of Lilly's political expenditures and lobbying activities to ensure that we fulfill our commitment to stewardship of corporate funds and risk minimization with respect to such activities. (Eli Lilly 2020)	0.929
<b>Executive Pay</b>	We did not accelerate the issuance of awards for our most senior executives. (Marriott 2021)	0.922
<b>Lawsuits</b>	According to the lawsuit, these damages arise from the Company's alleged failure to comply with safety and record maintenance regulations and false statements in public filings regarding the Company's safety practices. (Southwest 2021)	0.903
<b>Lobbying</b>	We are also concerned Lilly's lack of disclosure presents reputational risk when its lobbying contradicts company public positions. (Eli Lilly 2020)	0.920
<b>Tax Strategy</b>	Additionally, we repurchased \$10 million of shares during 2021 related to our tax withholding obligation associated with the vesting of employee restricted stock awards. (Marathon, 2021)	0.881
<b>Onboarding Policies</b>	We first introduce our culture to new crewmembers during the screening process and then at an extensive new hire orientation program at JetBlue University, our training center in Orlando. (JetBlue 2021)	0.595

# Airline Industry

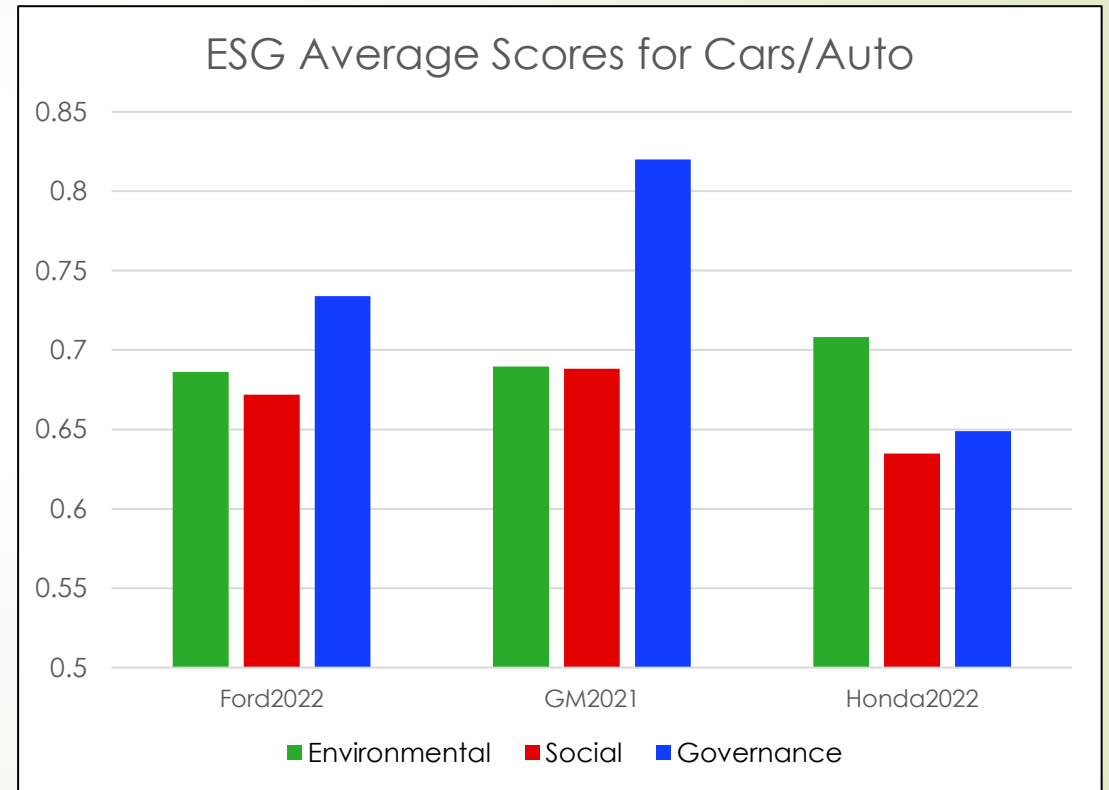
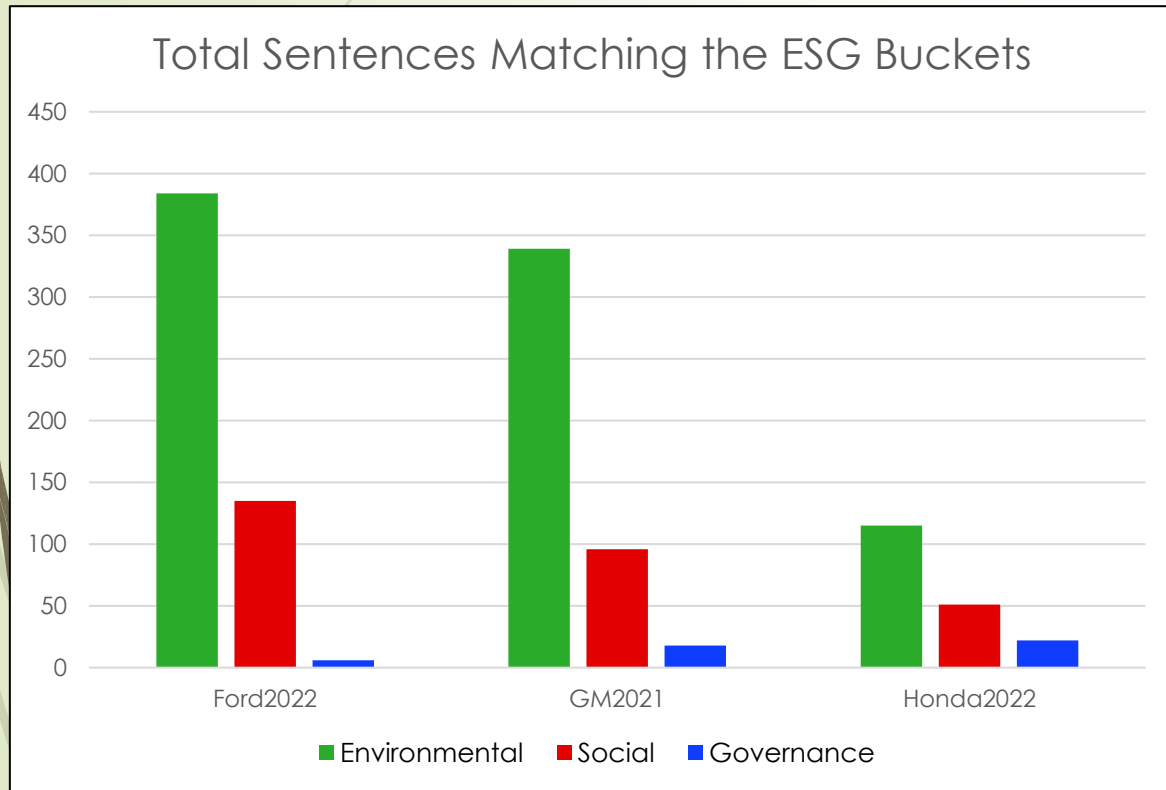


# Highest Scores: Southwest 2021 Annual Report

	<b>Environmental</b>	<b>Social</b>	<b>Governance</b>
<b>ESG Bucket</b>	Environmental Sustainability, 0.978	Employee Diversity , 0.776	Lawsuits, 0.931
<b>Sentence</b>	<i>"Environmental Sustainability: The Company remains steadfast in its desire to pursue, implement, and enhance initiatives to address the Company's impact on the environment."</i>	<i>"The Company's vast Employee training and development opportunities address, among other things, leadership development; diversity, equity, and inclusion; communication skills; and human trafficking awareness."</i>	<i>"Due to legal challenges to the vaccine mandate, the Company announced on December 20, 2021, that it is no longer targeting a January 4, 2022, deadline for compliance."</i>



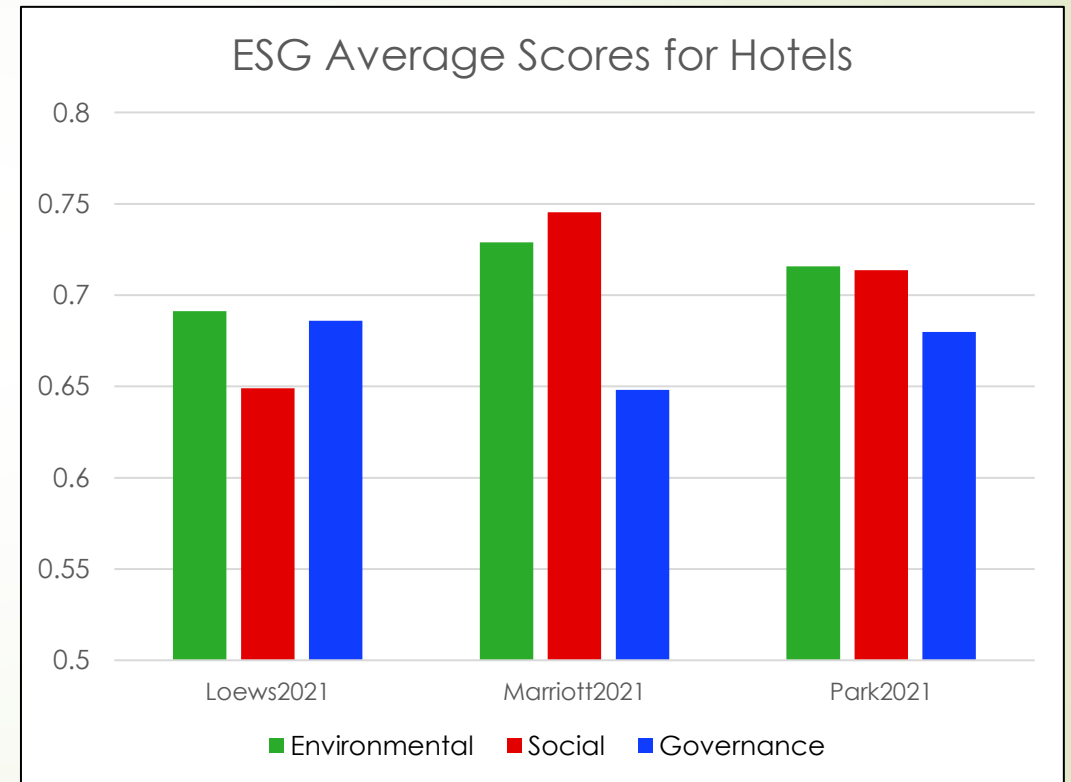
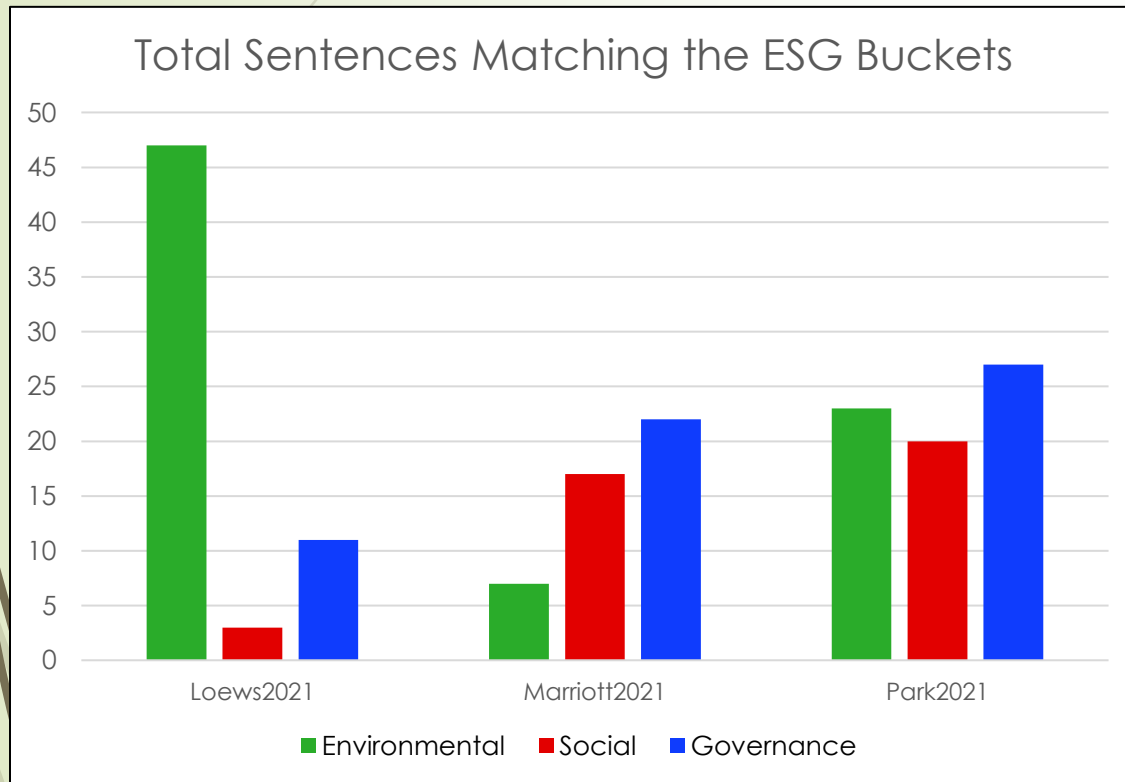
# Car/Automobile Industry



# Highest Scores: Ford 2022 Annual Report

	Environmental	Social	Governance
<b>ESG Bucket</b>	Air Quality, 0.964	Global Community Service, 0.984	Hiring Policies, 0.939
<b>Sentence</b>	<i>"These pollutants can affect air quality, particularly in urban areas, and potentially impact human health."</i>	<i>"You'll see the continued impact of Ford and the Ford Fund to strengthen communities around the world."</i>	<i>"Inclusive hiring is the foundation of what we believe and what we do."</i>

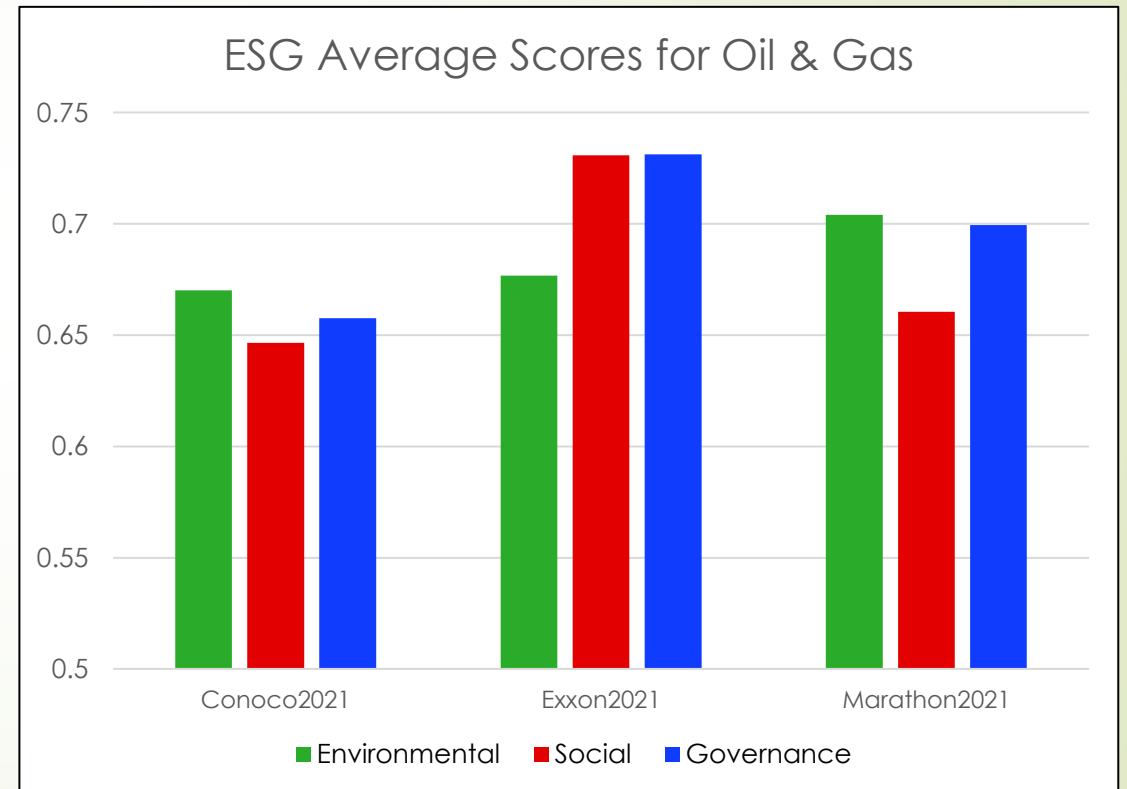
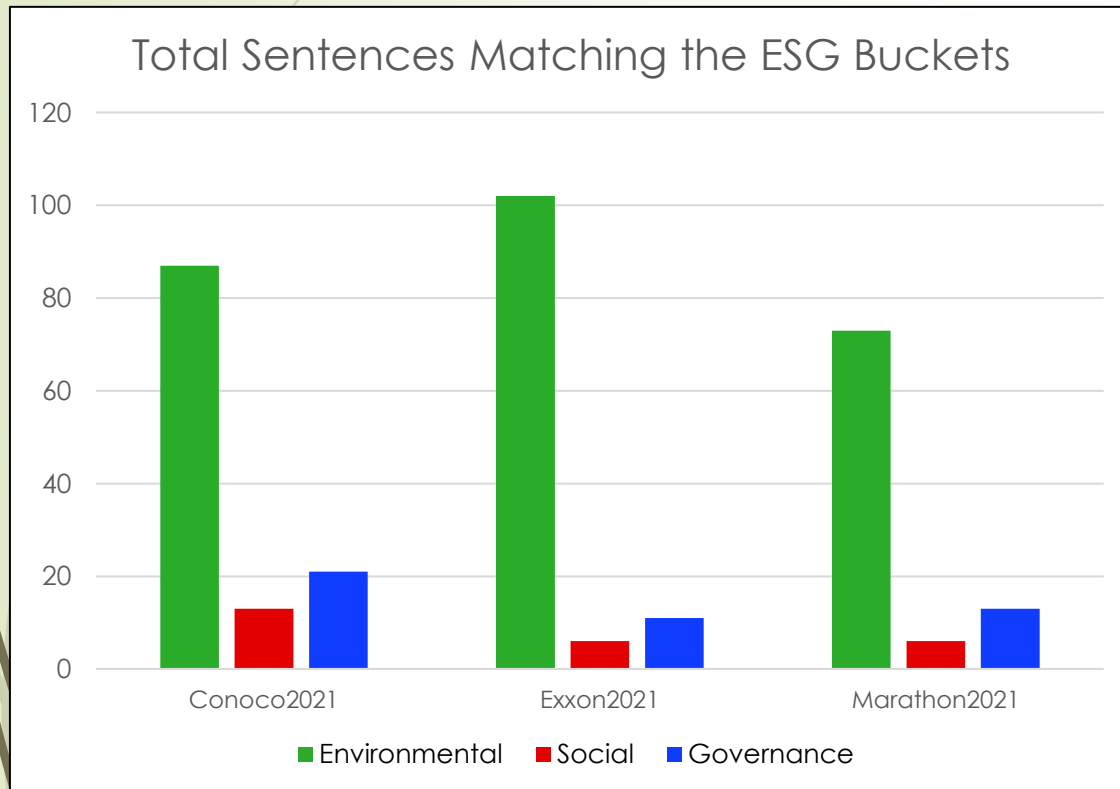
# Hotel/Hospitality Industry



# Highest Scores: Marriott 2021 Annual Report

	Environmental	Social	Governance
<b>ESG Bucket</b>	Climate Change, 0.905	Employee Health, 0.934	Executive Pay, 0.922
<b>Sentence</b>	<i>"Compliance with future climate-related legislation and regulation, and our efforts to achieve science-based emissions reduction targets, could be difficult and costly."</i>	<i>"For our U.S. property-level associates, we instituted a multipronged approach to encourage and support associate vaccination, including financial incentives, schedule flexibility, and health education."</i>	<i>"We did not accelerate the issuance of awards for our most senior executives."</i>

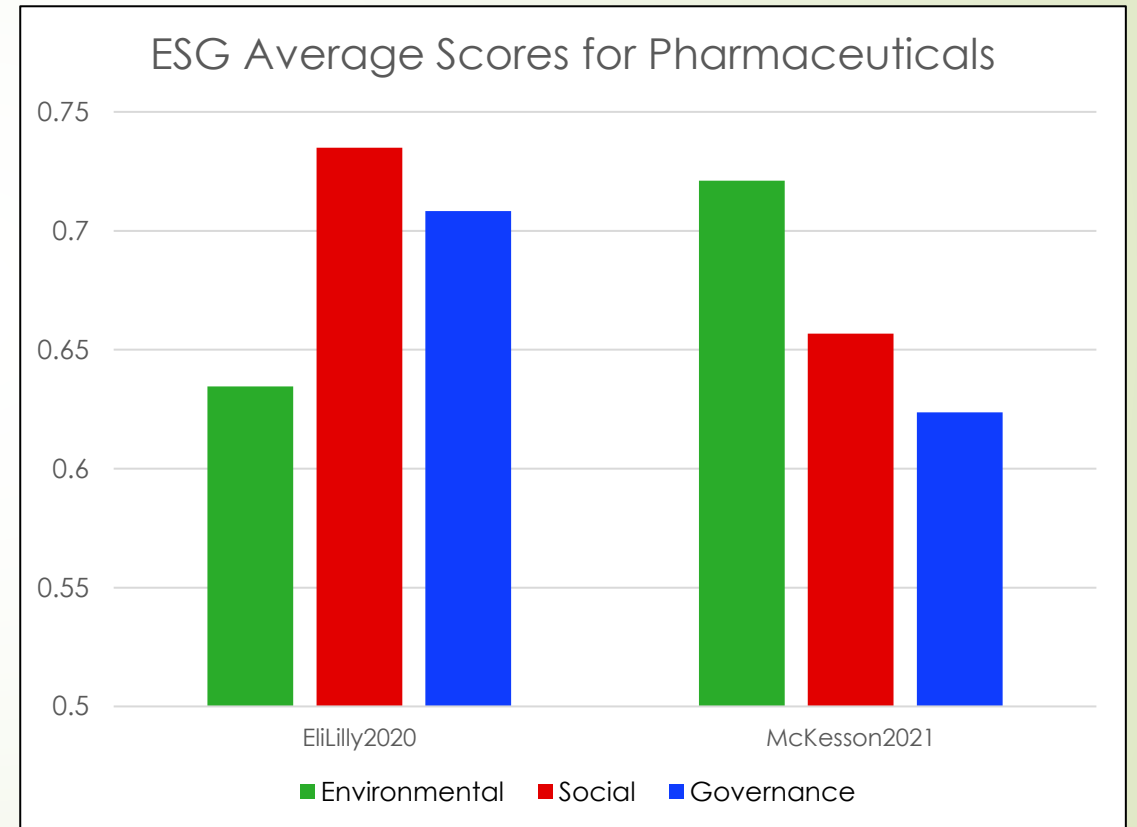
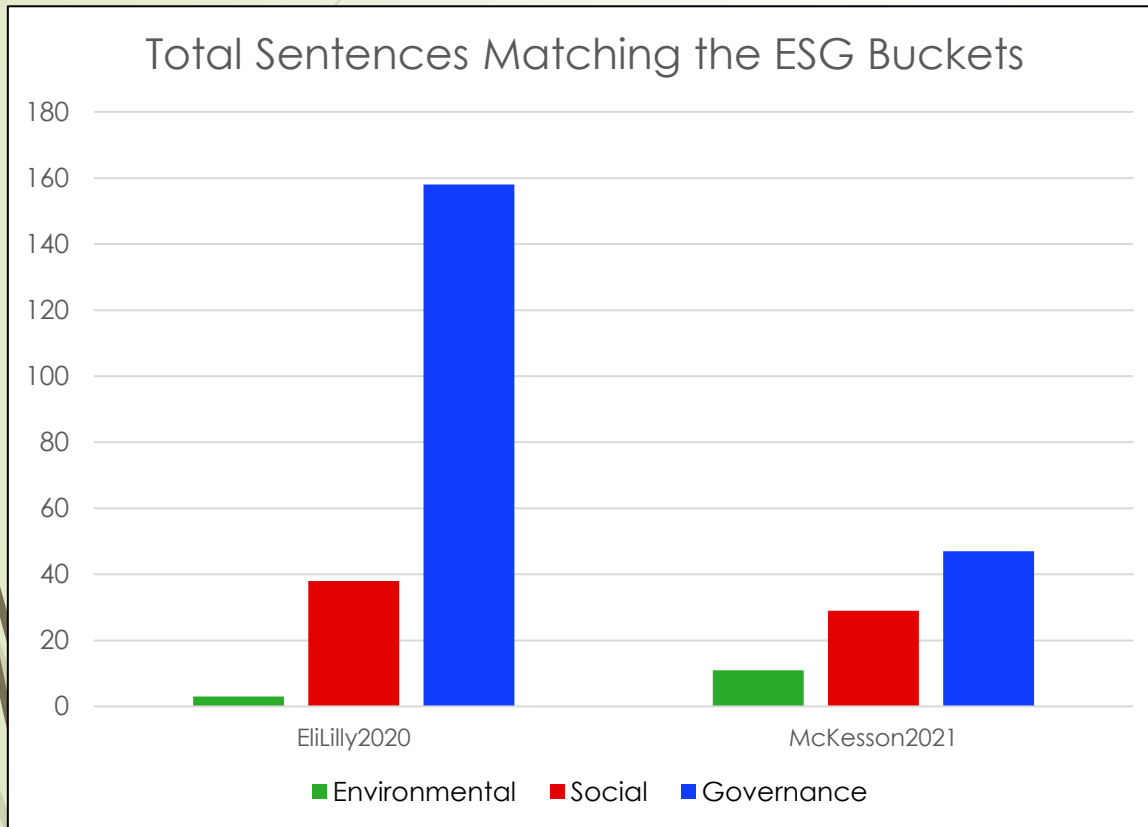
# Oil & Gas Industry



# Highest Scores: Exxon 2021 Annual Report

	Environmental	Social	Governance
<b>ESG Bucket</b>	Green Energy Initiatives, 0.944	Employee Diversity, 0.928	Executive Pay, 0.970
<b>Sentence</b>	<i>"Further opportunities include power and steam cogeneration and electrification of operations, using renewable or lower-emission power."</i>	<i>"With over 160 nationalities represented in the company, we encourage and respect diversity of thought, ideas, and perspective from our workforce."</i>	<i>"Directors, Executive Officers and Corporate Governance Executive Compensation Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters Certain Relationships and Related Transactions, and Director Independence Principal Accounting Fees and Services PART II PART III PART IV Exhibits, Financial Statement Schedules Form 10K Summary Item 15."</i>

# Pharmaceuticals Industry

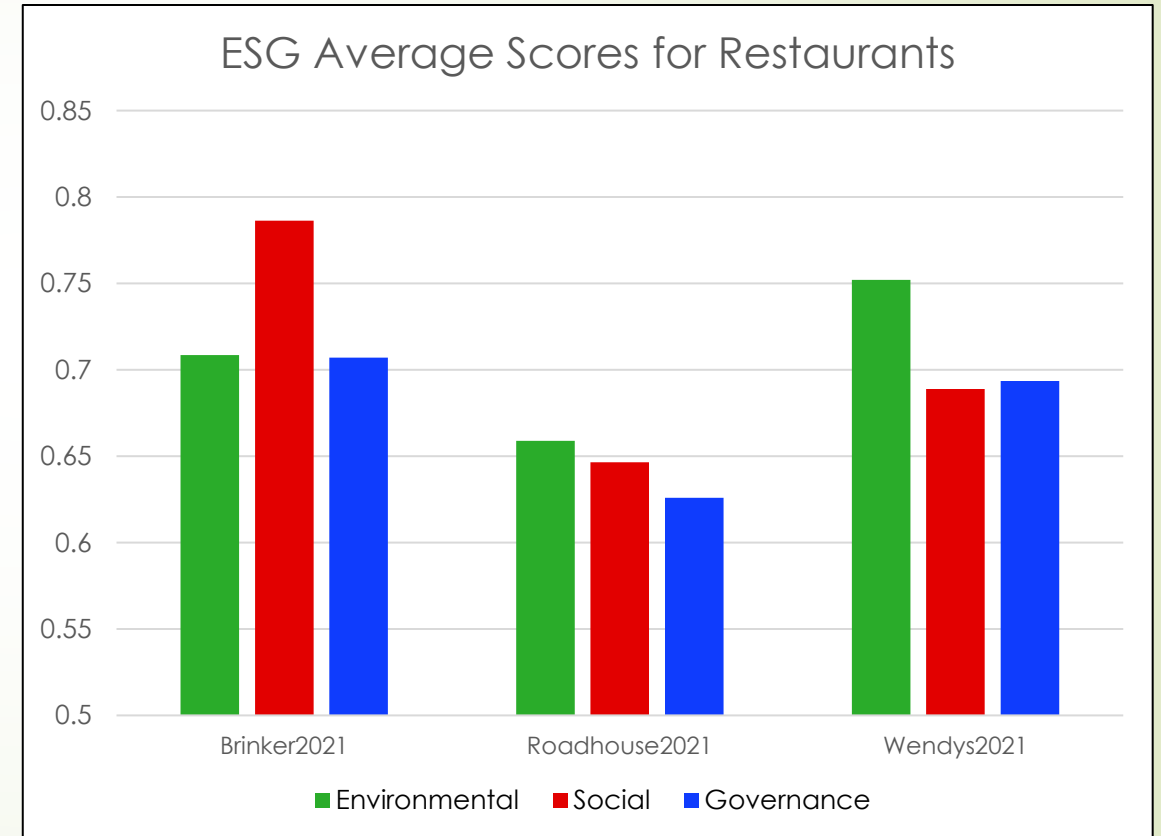
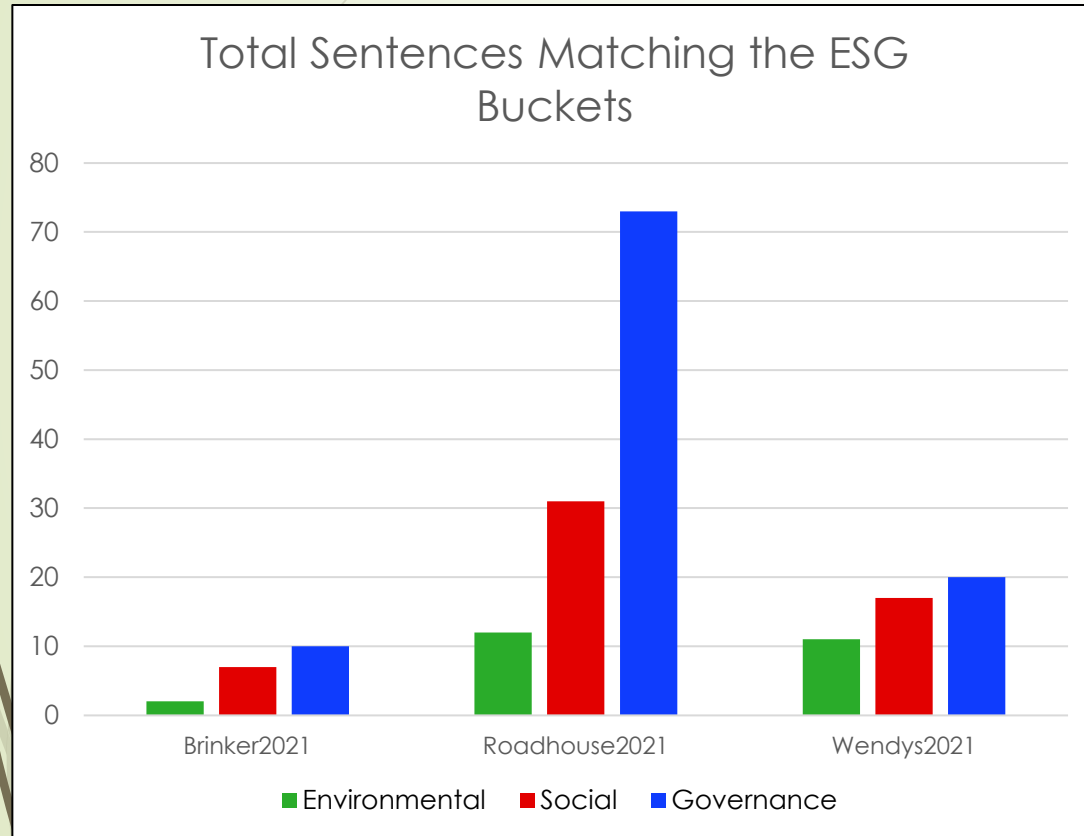


# Highest Scores: McKesson 2021 Annual Report

	Environmental	Social	Governance
<b>ESG Bucket</b>	Environmental Sustainability, 0.942	Employee Health, 0.882	Lawsuits, 0.885
<b>Sentence</b>	<i>"We meet this commitment through an environmental strategy and sustainability program."</i>	<i>"We expanded employee medical benefits covering COVID19 related visits, treatments, and testing as well as expanded telehealth options to protect employee safety."</i>	<i>"On August 8, 2018, the Company was served with a qui tam complaint pending in the United States District Court for the District of Massachusetts alleging that the Company violated the federal False Claims Act and various state false claims acts due to the alleged failure of the Company and other defendants to report providers who were engaged in diversion of controlled substances."</i>



# Restaurant Industry



# Highest Scores: Roadhouse 2021 Annual Report

	Environmental	Social	Governance
<b>ESG Bucket</b>	Environmental Sustainability, 0.972	Local Community Service, 0.965	Executive Pay, 0.944
<b>Sentence</b>	<i>"When it comes to conservation, bees, trees, and water are three resources we are passionate about protecting and preserving as a company."</i>	<i>"Our approach to media aligns with our focus on local store marketing and community involvement."</i>	<i>"It is our belief that a significant amount of each Named Executive Officer's compensation should be tied to the performance of the Company."</i>

## 28-Point Scoring System

Company	E	S	G	Total
Delta	5	6	4	<b>15</b>
Southwest	5	4	6	<b>15</b>
JetBlue	4	6	5	<b>15</b>
Ford	7	6	4	<b>17</b>
GM	6	6	6	<b>18</b>
Honda	7	6	4	<b>17</b>
Loews	6	3	6	<b>15</b>
Marriott	3	7	7	<b>17</b>
Park	5	6	6	<b>17</b>
Conoco	7	4	7	<b>18</b>
Exxon	5	4	4	<b>13</b>
Marathon	7	3	6	<b>16</b>
Eli Lilly	2	5	9	<b>16</b>
McKesson	3	6	6	<b>15</b>
Brinker	1	2	3	<b>5</b>
Roadhouse	3	6	6	<b>15</b>
Wendy's	2	6	6	<b>14</b>

# Analysis Summary

- Pros/Cons of each Scoring Metric
- Our Preferred Scoring Metric:
  - 28-point based system
  - Allows easy comparisons across companies and industries
  - Allows granular breakdown of differences across ESG buckets
- Notable/Interesting Findings

## Measure Commitment to Corporate ESG: Extracting from Annual Reports

► [https://bit.ly/mine\\_esg](https://bit.ly/mine_esg)



# What we learned

- ▶ Main ESG areas for industries and companies
- ▶ High score companies
- ▶ Main ESG areas overall