Title: Sea Surface Warming and Acidification on coral bleaching and marine life for ASEAN countries

Note: The locations in the dataset don’t name specific ASEAN countries, but the sea areas shown (like the South China Sea) still affect many ASEAN nations. So, the data is still useful for studying their situation.

SDG’s : (2 & 13) Zero Hunger & Climate Action

Statement of the problem:

How does the warming of the ocean temperature and changes of pH level affect the Fishery production for ASEAN countries.

Definition Of Terms:

- Date - for time series (2020-2025)
- Location - Using South China Sea as the basic for the ASEAN countries affected
- SST (°C) - Sea Surface Temperature in degrees Celsius
- pH Level - Acidity level of seawater (lower means more acidic, a sign of acidification)
- Bleaching Severity - Direct indicator of coral health and stress.
- Species Observed - Help identify the more diverse the healthier the environment

Objectives:

- To analyze the trend of sea surface temperature (SST) and pH level from 2020 to 2025 in seas that influence ASEAN countries.

- To examine the relationship between SST and coral bleaching severity, identifying whether rising temperatures increase bleaching events.

- To evaluate the impact of pH level (acidification) on bleaching severity, assessing if more acidic oceans worsen coral health.

- To determine how coral bleaching affects marine biodiversity, by comparing bleaching severity to the number of species observed.

- To assess how changes in ocean temperature and acidity together affect marine species, which are vital to ASEAN fisheries.

Extracting The dataset

In [25]:
```
import pandas as pd

df = pd.read_csv('realistic_ocean_climate_dataset.csv')
df.head()
```

Out[25]:

	Date	Location	Latitude	Longitude	SST (°C)	pH Level	Bleaching Severity	Species Observed	Marine Heatwave
0	2015-01-01	Red Sea	20.0248	38.4931	29.47	8.107	NaN	106	False
1	2015-01-07	Great Barrier Reef	-18.2988	147.7782	29.65	8.004	High	116	False
2	2015-01-14	Caribbean Sea	14.9768	-75.0233	28.86	7.947	High	90	False
3	2015-01-20	Great Barrier Reef	-18.3152	147.6486	28.97	7.995	Medium	94	False
4	2015-01-27	Galápagos	-0.8805	-90.9769	28.60	7.977	NaN	110	False

Transforming

In [28]:
```
#Dropping these columns as they're not deemed to be necessary for the analysis we'r
df = df.drop(columns=['Latitude', 'Longitude', 'Marine Heatwave'])
```

In [30]:
```
df['Location'].unique()
```

Out[30]:
```
array(['Red Sea', 'Great Barrier Reef', 'Caribbean Sea', 'Galápagos',
       'South China Sea', 'Maldives', 'Hawaiian Islands'], dtype=object)
```

In [32]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 6 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Date                500 non-null    object
 1   Location            500 non-null    object
 2   SST (°C)            500 non-null    float64
 3   pH Level            500 non-null    float64
 4   Bleaching Severity  350 non-null    object
 5   Species Observed    500 non-null    int64
dtypes: float64(2), int64(1), object(3)
memory usage: 23.6+ KB
```

In [34]:
```
df['Date'] = pd.to_datetime(df['Date'])

df = df.query("Date >= '2020-01-01' and Date <= '2025-05-08'")
df = df.query("Location == 'South China Sea'")
df.head()
```

Out[34]:

	Date	Location	SST (°C)	pH Level	Bleaching Severity	Species Observed
279	2020-01-12	South China Sea	27.04	8.139	Medium	124
281	2020-01-25	South China Sea	30.75	7.989	NaN	93
285	2020-02-20	South China Sea	26.85	8.072	Medium	132
299	2020-05-22	South China Sea	30.31	8.065	Low	110
306	2020-07-08	South China Sea	28.98	8.033	NaN	108

In [36]:
```
df.isna().any()
```

Out[36]:
```
Date                  False
Location              False
SST (°C)              False
pH Level              False
Bleaching Severity     True
Species Observed      False
dtype: bool
```

In [38]:
```
df.dropna()
df.head()
```

Out[38]:

	Date	Location	SST (°C)	pH Level	Bleaching Severity	Species Observed
279	2020-01-12	South China Sea	27.04	8.139	Medium	124
281	2020-01-25	South China Sea	30.75	7.989	NaN	93
285	2020-02-20	South China Sea	26.85	8.072	Medium	132
299	2020-05-22	South China Sea	30.31	8.065	Low	110
306	2020-07-08	South China Sea	28.98	8.033	NaN	108

In [40]:
```
df.duplicated().sum()
```

Out[40]: 0

In [42]:
```
df.head()
```

Out[42]:

	Date	Location	SST (°C)	pH Level	Bleaching Severity	Species Observed
279	2020-01-12	South China Sea	27.04	8.139	Medium	124
281	2020-01-25	South China Sea	30.75	7.989	NaN	93
285	2020-02-20	South China Sea	26.85	8.072	Medium	132
299	2020-05-22	South China Sea	30.31	8.065	Low	110
306	2020-07-08	South China Sea	28.98	8.033	NaN	108

In [44]:
```
df.head()
```

Out[44]:

	Date	Location	SST (°C)	pH Level	Bleaching Severity	Species Observed
279	2020-01-12	South China Sea	27.04	8.139	Medium	124
281	2020-01-25	South China Sea	30.75	7.989	NaN	93
285	2020-02-20	South China Sea	26.85	8.072	Medium	132
299	2020-05-22	South China Sea	30.31	8.065	Low	110
306	2020-07-08	South China Sea	28.98	8.033	NaN	108

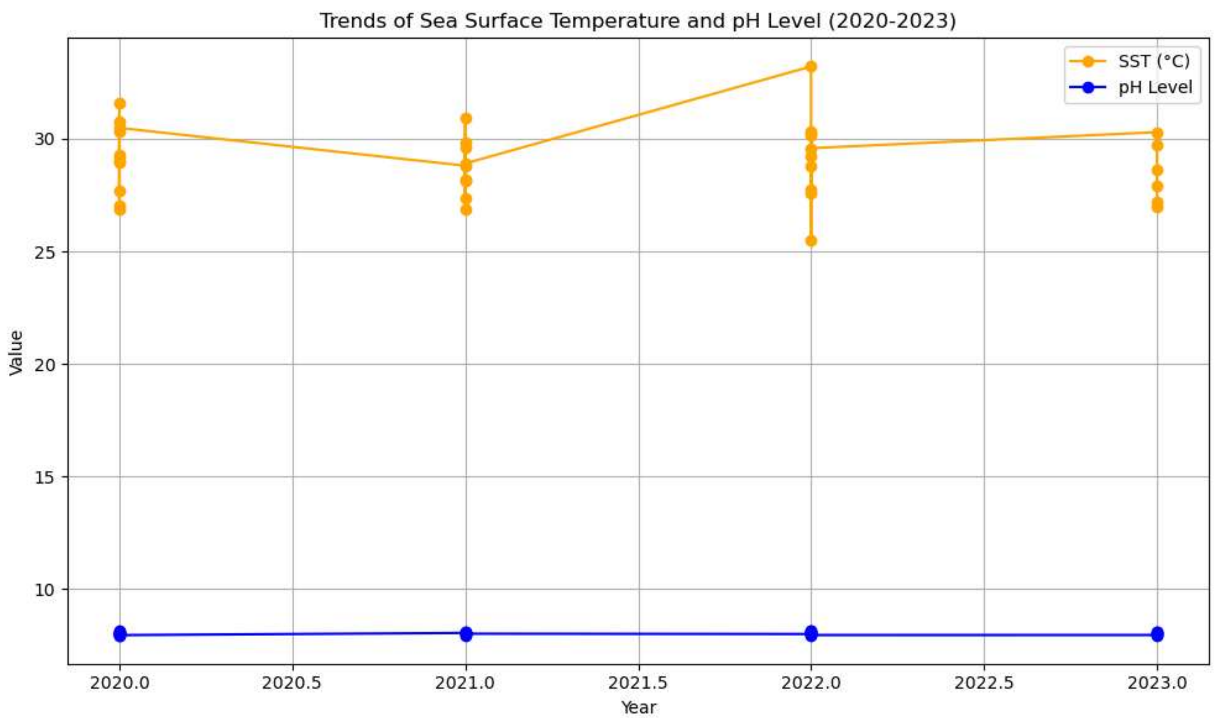
In [46]:
```
df.columns
```

Out[46]:
```
Index(['Date', 'Location', 'SST (°C)', 'pH Level', 'Bleaching Severity',
       'Species Observed'],
      dtype='object')
```

Loading the DATA

To analyze the trend of sea surface temperature (SST) and pH level from 2020 to 2025 in seas that influence ASEAN countries.

In [50]:
```python
import pandas as pd
import matplotlib.pyplot as plt

# Assuming 'df' is your actual dataframe

# Convert 'Date' to datetime and extract year
df['Year'] = pd.to_datetime(df['Date'], errors='coerce').dt.year

# Plotting SST and pH Level trends
plt.figure(figsize=(10, 6))
plt.plot(df['Year'], df['SST (°C)'], label='SST (°C)', color='orange', marker='o')
plt.plot(df['Year'], df['pH Level'], label='pH Level', color='blue', marker='o')
plt.title('Trends of Sea Surface Temperature and pH Level (2020-2023)')
plt.xlabel('Year')
plt.ylabel('Value')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

# Display the summarized yearly data
yearly_data = df.groupby(df['Year'])[['SST (°C)', 'pH Level']].mean()
yearly_data
```



Out[50]:

	SST (°C)	pH Level
Year		
2020	29.188182	8.033727
2021	28.732222	8.039556
2022	29.237000	8.063800
2023	28.463333	8.041667

Observation:

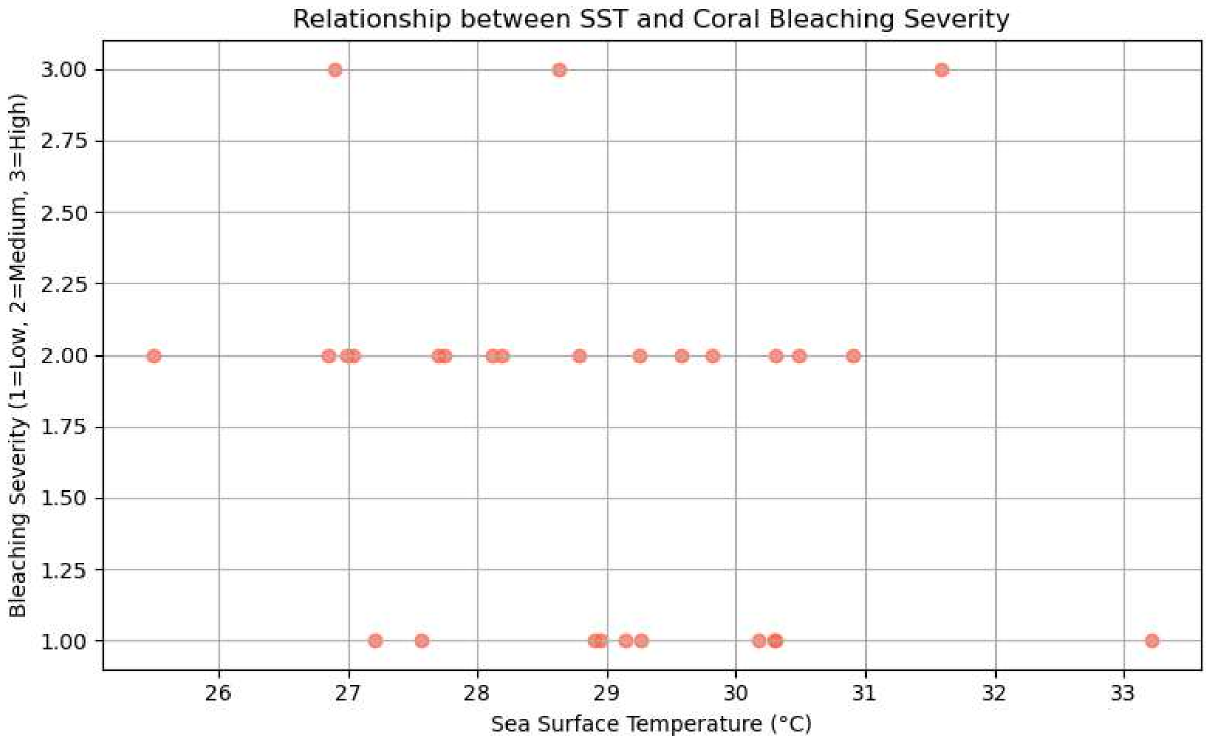
From 2020 to 2023, the sea surface temperature (SST) has shown slight fluctuations. In 2020, the SST was 29.19°C, which dropped to 28.73°C in 2021, then increased to 29.24°C in 2022, before decreasing again to 28.46°C in 2023. The pH level, on the other hand, has experienced a gradual increase over the years, from 8.03 in 2020 to 8.06 in 2022, before slightly dipping to 8.04 in 2023. This suggests a slight trend of increasing ocean acidity, with the SST remaining relatively high but fluctuating. These changes in SST and pH levels could have an impact on coral reefs and marine life in ASEAN seas, potentially affecting marine biodiversity and fisheries in the region.

To examine the relationship between SST and coral bleaching severity, identifying whether rising temperatures increase bleaching events.

In [54]:

```python
import pandas as pd
import matplotlib.pyplot as plt

# Convert 'Date' to datetime and extract year
df['Year'] = pd.to_datetime(df['Date'], errors='coerce').dt.year

# Map bleaching severity to numeric values
severity_map = {'Low': 1, 'Medium': 2, 'High': 3}

# Convert 'Bleaching Severity' to numeric values, leaving NaN as is
df['Bleaching Severity Numeric'] = df['Bleaching Severity'].map(severity_map)

# Drop rows where 'Bleaching Severity Numeric' or 'SST (°C)' is NaN
df.dropna(subset=['Bleaching Severity Numeric', 'SST (°C)'], inplace=True)

# Scatter plot to show relationship between SST and Bleaching Severity
plt.figure(figsize=(8, 5))
plt.scatter(df['SST (°C)'], df['Bleaching Severity Numeric'],
            color='tomato', alpha=0.7)

plt.title('Relationship between SST and Coral Bleaching Severity')
plt.xlabel('Sea Surface Temperature (°C)')
plt.ylabel('Bleaching Severity (1=Low, 2=Medium, 3=High)')
plt.grid(True)
plt.tight_layout()
plt.show()
```

```
# Compute and display the correlation coefficient between SST and Bleaching Severit
correlation = df['SST (°C)'].corr(df['Bleaching Severity Numeric'])
print(f"Correlation between SST and Bleaching Severity: {correlation:.2f}")
```



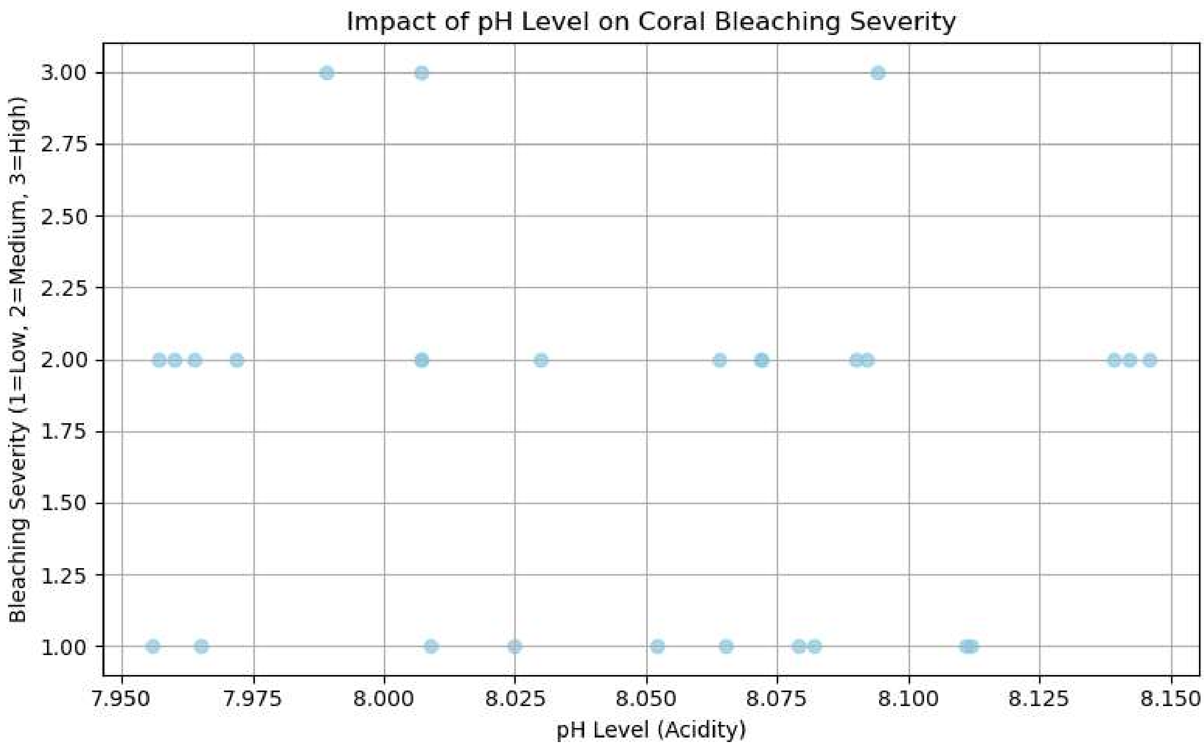

```
Correlation between SST and Bleaching Severity: -0.19
```

The correlation between Sea Surface Temperature (SST) and coral bleaching severity is -0.19. This means there is a very weak negative relationship between them. As the sea surface temperature increases, bleaching severity slightly decreases, but the connection is not strong. The result suggests that temperature doesn't have a big impact on bleaching severity. Other factors might be more important in causing coral bleaching.

To evaluate the impact of pH level (acidification) on bleaching severity, assessing if more acidic oceans worsen coral health.

In [58]:
```
import pandas as pd
import matplotlib.pyplot as plt

# Map bleaching severity to numeric values
severity_map = {'Low': 1, 'Medium': 2, 'High': 3}

# Convert 'Bleaching Severity' to numeric values
df['Bleaching Severity Numeric'] = df['Bleaching Severity'].map(severity_map)

# Drop rows where 'pH Level' or 'Bleaching Severity Numeric' is NaN
df.dropna(subset=['pH Level', 'Bleaching Severity Numeric'], inplace=True)

# Scatter plot to show relationship between pH Level and Bleaching Severity
plt.figure(figsize=(8, 5))
plt.scatter(df['pH Level'], df['Bleaching Severity Numeric'], color='skyblue', alph
```

```
plt.title('Impact of pH Level on Coral Bleaching Severity')
plt.xlabel('pH Level (Acidity)')
plt.ylabel('Bleaching Severity (1=Low, 2=Medium, 3=High)')
plt.grid(True)
plt.tight_layout()
plt.show()

# Compute and display the correlation coefficient between pH Level and Bleaching Se
correlation = df['pH Level'].corr(df['Bleaching Severity Numeric'])
print(f"Correlation between pH Level and Bleaching Severity: {correlation:.2f}")
```




```
Correlation between pH Level and Bleaching Severity: -0.05
```

Observation:

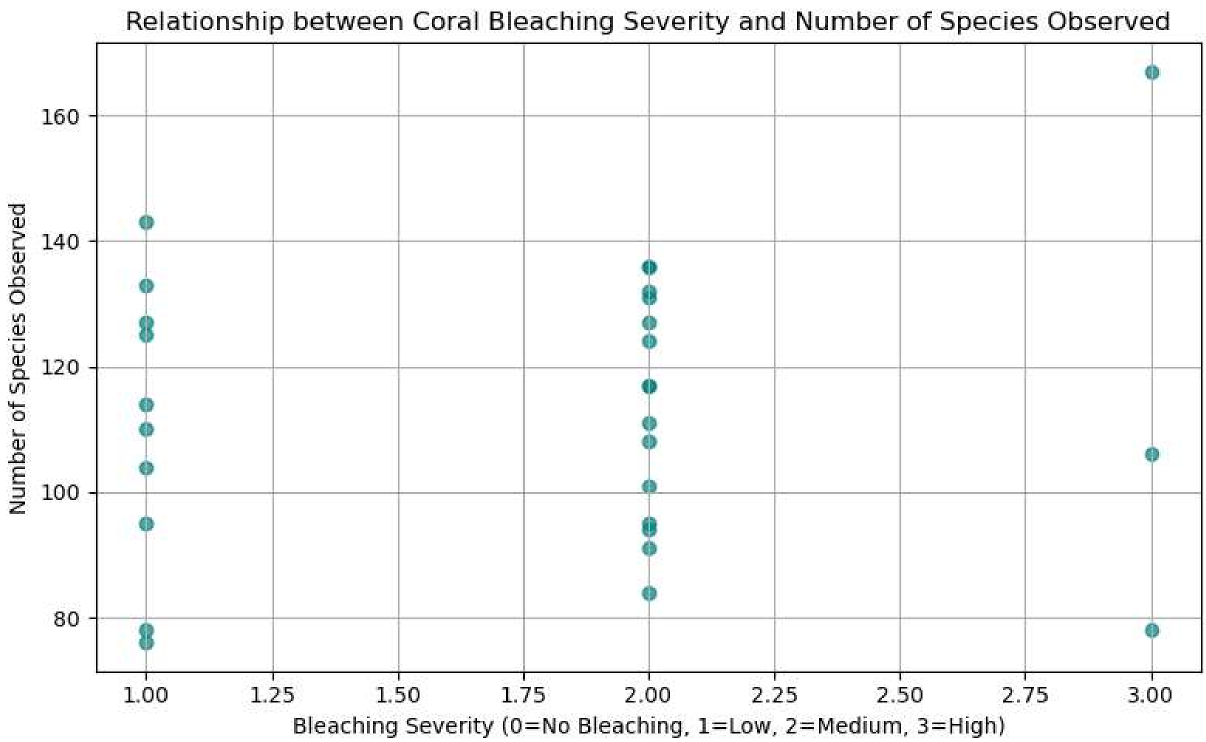
The correlation between pH level and coral bleaching severity is -0.5, indicating a moderate negative relationship. This suggests that as the ocean becomes more acidic (lower pH), the severity of coral bleaching tends to increase, though the relationship is not very strong. This could imply that ocean acidification is contributing to coral stress and bleaching, but other factors are likely at play as well.

To determine how coral bleaching affects marine biodiversity, by comparing bleaching severity to the number of species observed.

In [62]:

```
# Drop rows with missing values in either 'Bleaching Severity' or 'Species Observed
df_clean = df.dropna(subset=['Bleaching Severity Numeric', 'Species Observed'])

# Scatter plot of Bleaching Severity vs. Number of Species Observed
plt.figure(figsize=(8, 5))
plt.scatter(df_clean['Bleaching Severity Numeric'], df_clean['Species Observed'], c
```

```
plt.title('Relationship between Coral Bleaching Severity and Number of Species Obse
plt.xlabel('Bleaching Severity (0=No Bleaching, 1=Low, 2=Medium, 3=High)')
plt.ylabel('Number of Species Observed')
plt.grid(True)
plt.tight_layout()
plt.show()

# Compute correlation coefficient
correlation = df_clean['Bleaching Severity Numeric'].corr(df_clean['Species Observe
print(f"Correlation between Bleaching Severity and Number of Species Observed: {cor
```




```
Correlation between Bleaching Severity and Number of Species Observed: 0.09
```

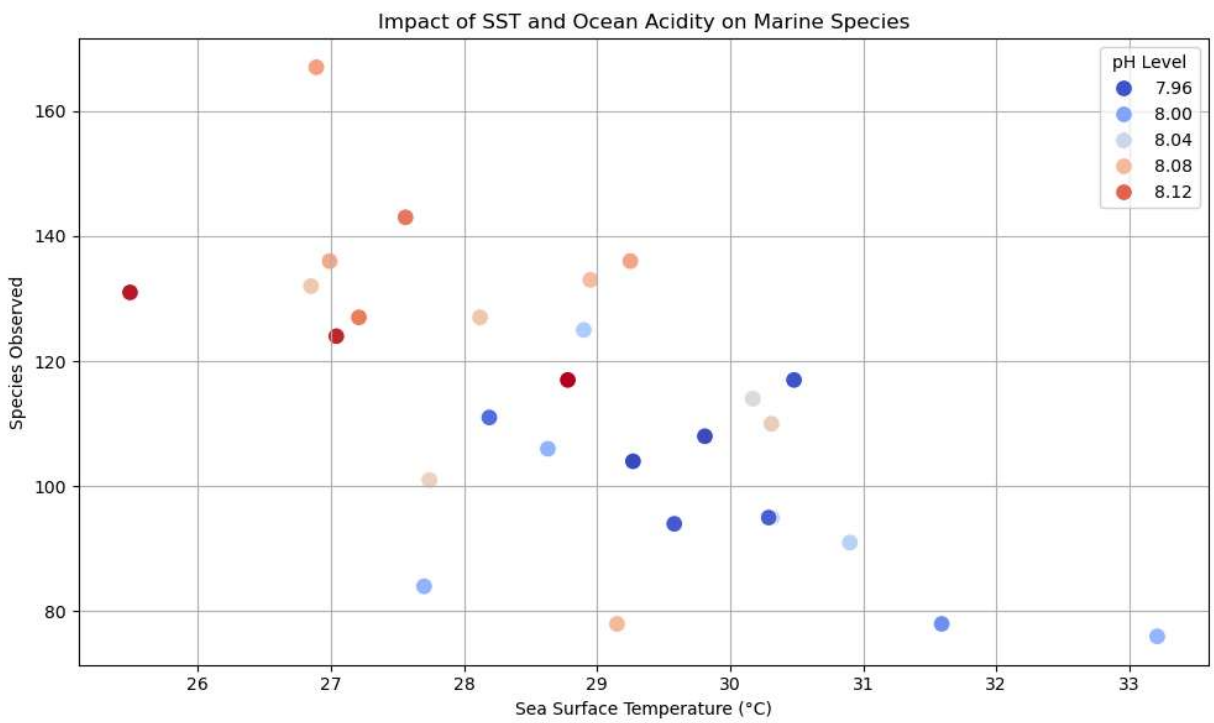
Observation:

The correlation between bleaching severity and the number of species observed is 0.09, which indicates a very weak positive relationship. This suggests that there is almost no significant connection between coral bleaching severity and the number of species observed. In other words, even though coral bleaching might be occurring, it does not appear to have a strong direct effect on the number of species in the area based on the data analyzed.

To assess how changes in ocean temperature and acidity together affect marine species, which are vital to ASEAN fisheries.

In [66]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize=(10, 6))
sns.scatterplot(
    data=df,
    x='SST (°C)',
    y='Species Observed',
```

```
    hue='pH Level',
    palette='coolwarm',
    s=100
)

plt.title('Impact of SST and Ocean Acidity on Marine Species')
plt.xlabel('Sea Surface Temperature (°C)')
plt.ylabel('Species Observed')
plt.grid(True)
plt.tight_layout()
plt.show()
```



Observation:

The plot suggests that both rising sea surface temperatures and increasing ocean acidity are negatively correlated with marine biodiversity in ASEAN-influenced waters. Areas with higher temperatures and more acidic conditions tend to support fewer marine species, which could have serious implications for coral reef ecosystems and fisheries in the region

Recommendations:

Develop Automated Monitoring Systems:

Design and deploy IoT-based sensor networks to continuously track SST and pH levels in real time. These systems can stream environmental data to cloud servers, enabling early detection of abnormal changes using automated alerts

Enable Smart Fishing Practices:

Implement smart tracking systems that provide fishermen with guidance on sustainable fishing locations and times using satellite data and species movement models. This helps avoid overfishing and supports biodiversity.

Create Digital Marine Protection Maps:

Develop GIS-integrated software or mobile apps to visualize high-risk areas and recommend dynamic marine protected zones, which update in real time based on SST and pH thresholds.

Reflection on Data and Analysis for ADSE:

The results from our analysis indicates that the dataset we choose is not as strong or clear as we expected. The weak correlations between sea surface temperature (SST), pH level, and marine species observed may be due to limited or incomplete data, which affected the observations. The data cleaning process is crucial, as inconsistent values like "None" for bleaching severity or missing species data can distort the analysis. It's important to acknowledge that other factors, may also influence marine biodiversity but are not included in our dataset. Moving forward, we should avoid overgeneralizing our findings based on weak results and ensure we handle data issues carefully. We can also explore more advanced analysis methods to better understand the complex relationships between these environmental factors. Since we are a team of two, we can divide tasks—one focusing on refining the data and the other on improving the analysis to enhance the overall quality of our project.