**SDG – 14: LIFE BELOW WATER**

Sustainable Marine Resource Management: Statistical Insights and Policy Recommendations for India's Exclusive Economic Zone

## AS1117: PROBABILITY AND STATISTICS

FACULTY GUIDE

Dr. Umesh Gupta

Dr. Rabia Kamra

PREPARED BY: Group-5 (Section - B)

INDRAJIT ROY (2023BTECH037)

SARVESH GULGULIA (2023BTECH073)

SHUBHAM SHARMA (2023BTECH081)

SNEHA MISHRA (2023BTECH083)



Institute of Engineering and Technology (IET)

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Content** | **Page No.** |
| 1. | Abstract | 3 |
| 2. | Introduction | 4-6 |
| 3 | GOAL | 7-9 |
| 4 | Target | 10 |
| 5 | Methodology | 11-15 |
| 6 | Problem Statements | 16-25 |
| 7 | Tables | 26-30 |
| 8 | Codes | 31-55 |
| 9 | Graphs | 56-72 |
| 10 | Conclusion | 73 |

**ABSTRACT**

The conservation and sustainable use of oceans and their resources is a critical goal in addressing the challenges posed by overfishing, habitat destruction, and climate change. The United Nations Convention on the Law of the Sea (UNCLOS) serves as the foundational legal framework guiding nations in the responsible management of marine environments and resources. As highlighted in paragraph 158 of "The Future We Want," a document emerging from the United Nations Conference on Sustainable Development, the effective implementation of international law is essential for ensuring the health of marine ecosystems and the livelihoods dependent on them. By adhering to UNCLOS, countries can enhance collaboration, enforce environmental protections, and foster sustainable practices, ultimately promoting a balanced and resilient ocean ecosystem for future generations.

**INTRODUCTION**



Fig 1: Sustainable Development Goals

Sustainable Development is described as the process of meeting the needs that are essential to the present generation without compromising the capability for meeting the needs when the future generation requires them.

The SDGs were, therefore, a product of 2015 by the UNGA through the Post-2015 Development Agenda with the sole aim of formulating a universal framework for future growth. This was a movement since the MDGs expired in 2015. It was formally adopted by the UNGA Resolution through a document titled "2030 Agenda for Sustainable Development," but more commonly referred to as Agenda 2030.

There are 17 Sustainable Development Goals, each having its prime focus:

• No Poverty: This involves ensuring that all people in the country have enough to eat and adequate clothing and housing to meet their needs, with no one having to struggle in extreme poverty.

• Zero Hunger: This is the attainment of better health by reducing food shortage and malnutrition.

• Good Health and Well-Being: Adequate nutrition and a pollution-free environment contribute to a reduced rate of health disorders and mortality.

• Quality Education: Setting up well-equipped schools and employing qualified teachers to ensure education is available for all, raising awareness of its benefits.

• Clean Water and Sanitation: Ensuring clean water is available along with proper sanitation facilities to support hygiene and health.

• Affordable and Clean Energy: Improving access to energy for all while conserving and protecting resources for future use.

• Decent Work and Economic Growth: Employment opportunities under decent conditions help solve unemployment problems.

• Industry, Innovation, and Infrastructure: Building industries for job creation to ensure workers receive fair wages.

• Reduced Inequalities: Ensuring nondiscrimination in policies, institutions, and society based on gender, caste, color, or origin.

• Sustainable Cities and Communities: Focusing on making human settlements safe, equitable, and providing access to opportunities in urban areas.

• Sustainable Consumption and Production: Achieving an appropriate supply and demand balance toward sustainable production and consumption.

• Climate Action: Developing preemptive measures to address natural calamities and climate change.

• Life Below Water: Conserving oceans and seas through proper management of marine resources to ensure availability in the future.

• Life on Land: Combating deforestation and addressing land degradation caused by population growth to restore balance.

• Peace, Justice, and Strong Institutions: Establishing justice systems respecting the rule of law and creating fair, order-keeping institutions.

• Partnership for the Goals: Promoting international cooperation among countries toward sustainable development goals.

Goal 14: Life Below Water



Fig 2: Life below Water(SDG Goal 14)

SDG 14 focuses on conserving and sustainably using the oceans, seas, and marine resources for sustainable development. This goal is vital because oceans are essential to global ecosystems, providing food, regulating the climate, and supporting biodiversity. However, marine environments face significant threats, including pollution, overfishing, and climate change, which undermine their health and sustainability. By addressing these challenges, SDG 14 aims to protect marine ecosystems, ensure the livelihoods of communities dependent on oceans, and preserve marine resources for future generations. Key areas of focus include reducing marine pollution, protecting marine biodiversity, promoting sustainable fishing practices, and increasing scientific knowledge to better manage ocean resources. Achieving this goal is crucial for creating a more balanced and sustainable planet.

**Why is Life Below Water important?**

Oceans are critical to life on Earth for several reasons:

1. **Regulating the climate:** Oceans absorb a significant portion of carbon dioxide emissions and heat, helping to mitigate climate change and regulate global temperatures.
2. **Supporting biodiversity:** Marine ecosystems are home to an immense variety of species that play essential roles in maintaining ecological balance.
3. **Providing livelihoods:** Millions of people rely on oceans for their livelihood through fishing, aquaculture, and tourism industries.
4. **Ensuring food security:** Oceans are a vital source of protein for billions of people, particularly in developing countries.

**How does the decline of marine ecosystems affect people and the planet?**

1. **Food insecurity:** Overfishing and habitat destruction reduce fish populations, jeopardizing food supplies for communities that depend on marine resources.
2. **Economic losses:** Industries such as fisheries, tourism, and coastal economies suffer when marine ecosystems are degraded.
3. **Biodiversity loss:** The destruction of marine habitats leads to a decline in species, disrupting ecosystems and threatening global ecological balance.
4. **Increased climate vulnerability:** Degraded oceans are less effective at absorbing carbon dioxide, exacerbating the impacts of climate change.

**Why should Life Below Water matter to everyone?**

1. **Preserving resources for future generations:** Protecting marine ecosystems ensures that oceans continue to provide food, livelihoods, and ecological services.
2. **Mitigating climate change:** Healthy oceans play a crucial role in regulating the Earth’s climate by absorbing carbon emissions and maintaining weather patterns.
3. **Supporting sustainable development:** By conserving marine resources, we can foster economic growth and social well-being for coastal and island communities.
4. **Safeguarding biodiversity:** Protecting marine life ensures the stability and resilience of ecosystems that benefit all life on Earth.

Achieving SDG 14 is essential for maintaining a sustainable balance between human activities and the health of the oceans, ensuring a thriving planet for current and future generations.

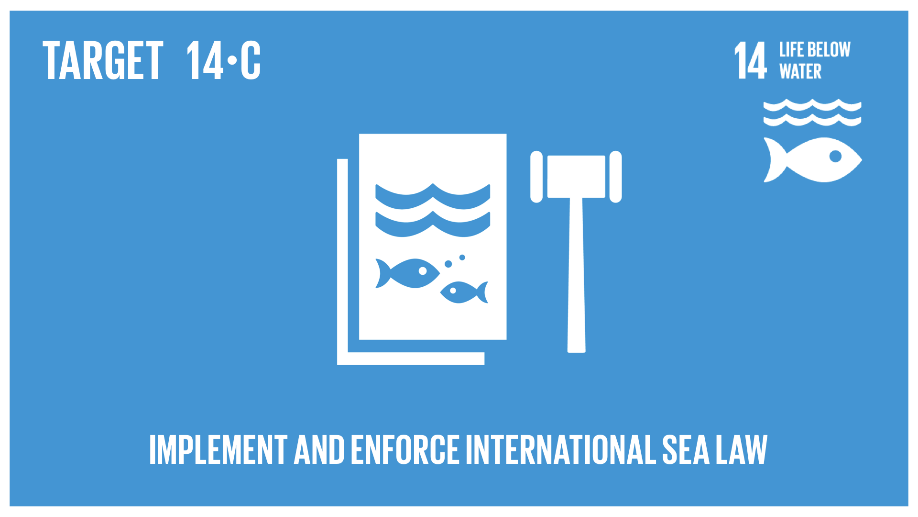


Fig 3: Target 14.c

**TARGET 14.C**

Enhance the conservation and sustainable use of oceans and their resources by implementing international law as reflected in United Nations Convention on the Law of the Sea, which provides the legal framework for the conservation and sustainable use of oceans and their resources, as recalled in paragraph 158 of "The future we want".

**Methodology**

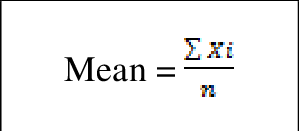
**Following Methods are used in the report:**

**1. Descriptive Statistics**

Descriptive statistics focus on summarizing and describing the essential features of a dataset.

**Key Metrics and Formulas:**

* **Mean (Arithmetic Average):**



Where Xi are data points, and N is the total number of observations.

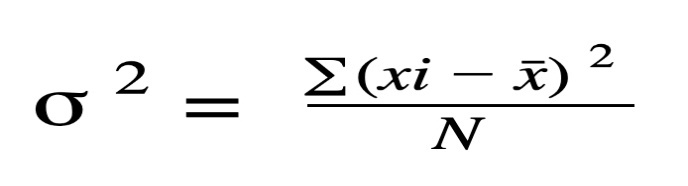
 Median**:**  
The middle value when data points are arranged in ascending order.

 Mode**:**  
The most frequently occurring value in the dataset.

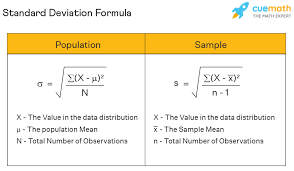
 Range**:**

Range=Maximum Value−Minimum Value

**Variance (Sample):**  
Measures how data points spread around the mean.



**Standard Deviation:**  
The square root of variance, indicating average deviation from the mean.



**Skewness:**

Measures asymmetry in the distribution of data.

**Kurtosis:**  
Describes the "tailedness" of the data distribution.

**2. Inferential Statistics**

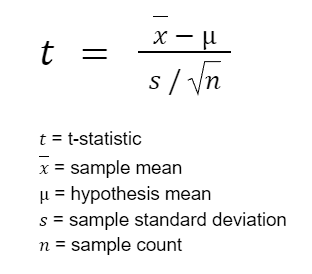
Inferential statistics involve drawing conclusions about a population based on sample data. Key techniques include hypothesis testing and confidence intervals.

T-Tests

(a) One-Sample T-Test:

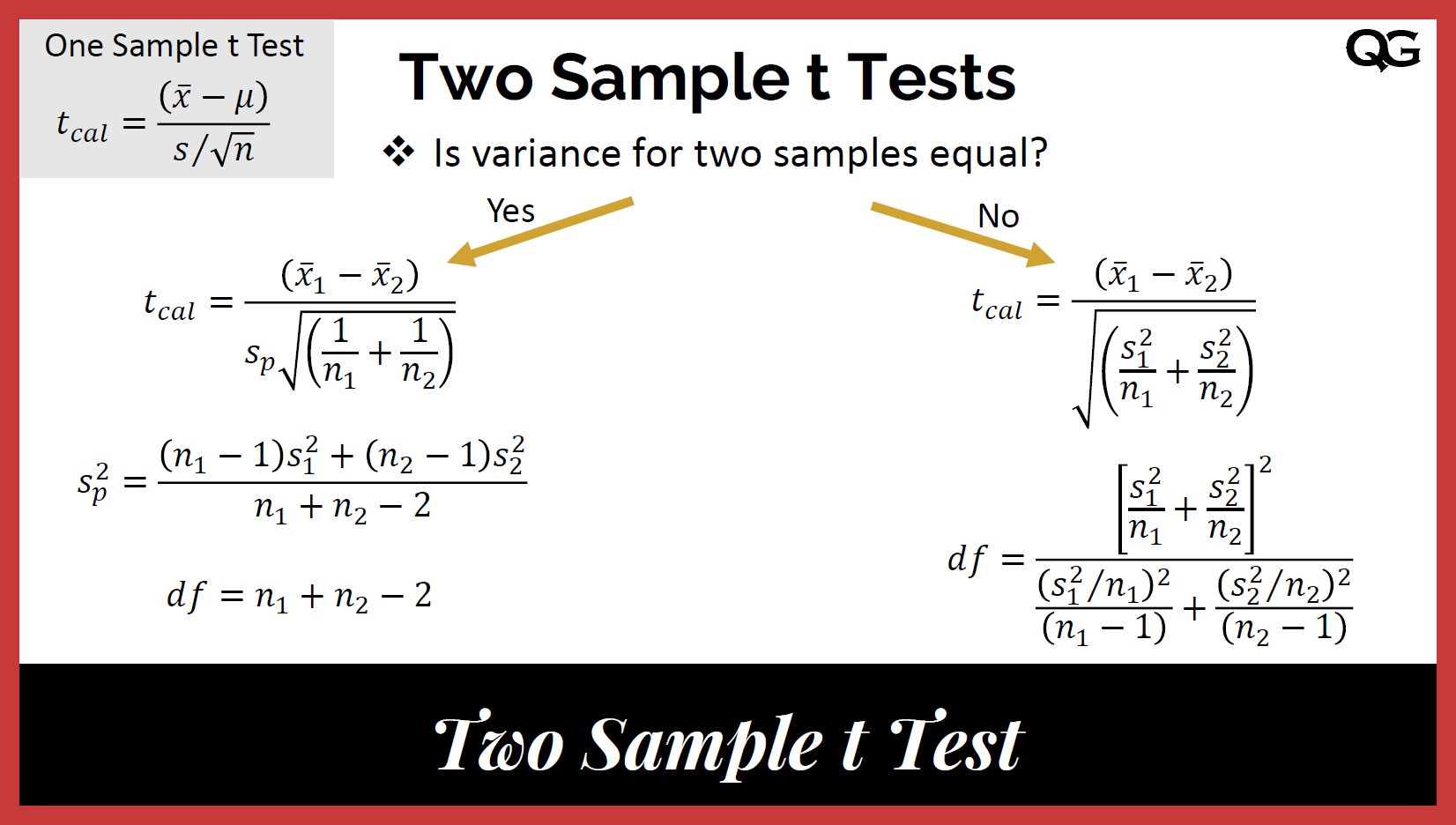
Used to determine if a sample mean differs significantly from a known population mean.

Formula:



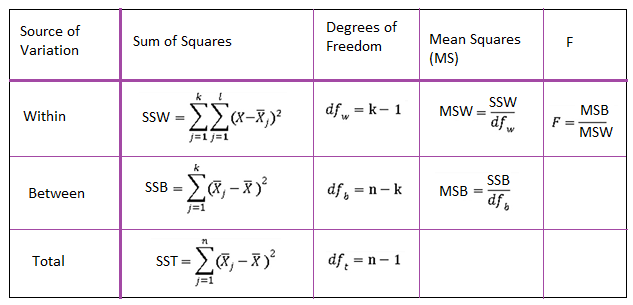
**(b) Independent Two-Sample T-Test:**

Compares the means of two independent groups.



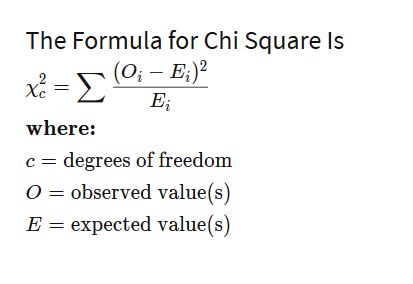
**3. Analysis of Variance (ANOVA)**

ANOVA tests if the means of three or more groups differ significantly.



**4. Chi-Square Test**

Used to test relationships between categorical variables.



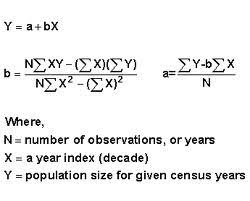
Applications include:

* **Goodness-of-Fit Test:** Checks if data fits a particular distribution.
* **Test of Independence:** Examines the relationship between two categorical variables.

**5. Regression Analysis**

Regression evaluates relationships between a dependent variable and one or more independent variables.

**Simple Linear Regression Formula:**



**7. Hypothesis Testing**

Tests whether a claim about a population parameter is true.

**Key Steps:**

1. **State Hypotheses:**
   * Null Hypothesis (H0​): Assumes no effect or no difference.
   * Alternative Hypothesis (Ha​): Indicates presence of an effect or difference.
2. **Choose Significance Level (α\alphaα):** Typically, α=0.05.
3. **Compute Test Statistic:** Depending on the test
4. **Compare to Critical Value or p-value:** Decide whether to reject

**Problem Statements**

Problem Statement 1: Fishery and stock dynamics of ribbonfish Trichiurus lepturus in the Indian Exclusive Economic Zone

* Q1 Is there a significant difference in the mean annual landings between years ?
* Q2. Is there an association between the gear type (trawl net or gillnet) and the state?
* Q3. Does the fishing gear type (Trawl net vs. Gillnet) significantly affect the seasonal landings of T. lepturus?

if the observed variations in landings are statistically significant and if the factors interact in a meaningful way

Problem Statement 2: Financial and logistical constraints prevent India from effectively enforcing UNCLOS pollution control measures in the Exclusive Economic Zone.

* Q1 How do Annual Pollution Incidents relate to the level of Coastal Enforcement Personnel over the years? Does increasing personnel correlate with fewer pollution incidents?

Descriptive statistics were calculated for Annual Pollution Incidents and Coastal Enforcement Personnel, including mean, median, standard deviation, range, kurtosis, and skewness.

* Q2. What is the relationship between the Marine Pollution Budget Ratio (Percent) and Annual Pollution Incidents, and how does personnel availability per kilometre of coast influence the number of pollution incidents?

The correlation between Training Frequency and Annual Pollution Incidents was analyzed using Pearson's correlation coefficient. The results were visualized using a scatter plot.

Correlation between Training Frequency and Annual Pollution Incidents: 0.2219846746552922, p-value: 0.33348878627204254

The correlation between Training Frequency and Annual Pollution Incidents is 0.22, with a p-value of 0.33. This indicates a weak positive correlation that is not statistically significant (p-value > 0.05). Therefore, changes in training frequency do not strongly influence the number of pollution incidents.

* Q3. Which factors—Marine Pollution Budget Ratio, Personnel per km of Coast, and Pollution Response Vessels—have the strongest influence on the number of Annual Pollution Incidents, and how can these insights guide resource allocation?

Regression analysis was performed to identify factors influencing Annual Pollution Incidents. Both linear and polynomial regression models were used to capture the relationships and predict future trends.

Linear Regression

The linear regression model provides the following coefficients and intercept:

Coefficients:

Marine Pollution Budget Ratio: -65.62

Personnel per km of Coast: 354.66

Pollution Response Vessels: 2.66

Intercept: 1207.21

These coefficients indicate that an increase in Personnel per km of Coast is associated with an increase in Annual Pollution Incidents, while an increase in Marine Pollution Budget Ratio is associated with a decrease in incidents. The impact of Pollution Response Vessels is minimal.

Polynomial Regression

The polynomial regression model provides the following coefficients and intercept:

Coefficients:

Intercept: 0.00

Marine Pollution Budget Ratio: 9.60

Personnel per km of Coast: 2062.23

Pollution Response Vessels: -8.07

Interaction terms and higher-order terms: Various values

Intercept: -58.32

The polynomial regression model captures more complex relationships between the variables and provides a more accurate prediction of future trends.

Future Trends Prediction

The polynomial regression model predicts a gradual decrease in Annual Pollution Incidents over the next decade, with the number of incidents dropping from 1327.96 in 2025 to 438.37 in 2034.

* Q4. Does training frequency per year significantly impact the number of Annual Pollution Incidents, and how does this vary with prosecution rates?

Hypothesis

NULL Hypothesis: There is no impact of training frequency on annual pollution incidents.

Alternate Hypothesis: There is impact of training frequency on annual pollution incidents.

ANOVA result: F-statistic=91.98228944600301, p-value=6.35472853811979e-12

The ANOVA result indicates that training frequency per year does not significantly impact Annual Pollution Incidents (p-value > 0.05). This suggests that variations in training frequency do not lead to significant changes in the number of pollution incidents.

Problem Statement 3

Loss of marine biodiversity

Q1Is there a significant difference in marine biodiversity between regions with low and high freshwater runoff?

The t-test compares the biodiversity in two different groups (low vs. high runoff) to see if they differ significantly.

Q2Do marine biodiversity levels differ significantly across regions with low, medium, and high levels of human impact?

ANOVA compares the biodiversity between three or more groups (low, medium, and high impact) to determine if at least one group differs from the others.

Q3How does the level of freshwater withdrawal predict the decline in marine species count?

Linear regression looks for a relationship between one independent variable and a dependent variable It helps quantify how changes in freshwater withdrawal might impact biodiversity.(Independent variable: Freshwater withdrawal percentage in regions (measured in %).

Dependent variable: Marine species count )

**PROBLEM STATEMENT 4**:

"Identifying sustainable fishing gear usage patterns across different regions to inform policies and practices that promote environmental protection, economic viability, and social equity in the fisheries sector."

"Optimizing the use of fishing gears to balance ecological, economic, and social objectives in coastal communities."

1.Understanding the current usage patterns and relationships between different fishing gear types.

2. Identifying ways to optimize the use of these gears to achieve a balance between environmental, economic, and social considerations in coastal regions where fishing is a critical activity.

**QUESTIONS:**

**QUES 1. How the use of each gear type has changed over time.**

**QUES 2. How the gear usage differs between the two regions.**

To analyze how the use of each fishing gear type has changed over time and how the gear usage differs between regions, I will perform a descriptive analysis of the data shown in the line graph and scatter plot.

The line graph provides a clear visualization of the trends in fishing gear contribution percentages from 2012 to 2017 for the three regions - EAS, WBoB, and India. By examining the line plots for each gear type, we can observe how their usage has changed over time within each region.

For example, in the EAS region, the trawl net contribution has remained relatively stable around 40-45%, while the gillnet usage has increased from 25% to 35% over the time period. In contrast, the WBoB region has seen a decrease in trawl net usage from 30% to 20%, accompanied by a more consistent gillnet contribution of 30-35%.

The scatter plot further highlights the differences in gear usage patterns between the regions. We can see that the data points for each region form distinct clusters, indicating that the relative prevalence of the various gear types varies significantly across EAS, WBoB, and India.

To quantify these regional differences, we could perform statistical tests like ANOVA (Analysis of Variance) to determine if the mean contributions of the gear types are significantly different between the regions. Additionally, t-tests could be used to identify which specific gear types have statistically significant differences in their usage across the regions.

In summary, the descriptive analysis of the line graph and scatter plot suggests that the usage of fishing gear types has evolved differently over time within each region, and there are also notable differences in the gear composition between the regions. Further statistical analysis could provide more robust insights into these trends and regional variations.

**QUES 3. Are the dominant fishing gears in each region associated with sustainable fishing practices?**

**How do the contributions of mechanized (e.g., trawlnets) and non-mechanized gears compare?**

To assess whether the dominant fishing gears in each region are associated with sustainable fishing practices, and to compare the contributions of mechanized and non-mechanized gears, I will perform a correlation analysis using the data provided in the correlation matrix.

The correlation matrix shows the strength and direction of the relationships between the usage of various fishing gear types. By examining the correlation coefficients, we can gain insights into how the different gears are interconnected.

CONCLUSION:

1. Temporal Trends in Gear Usage:
   1. The line graph analysis shows how the usage of different fishing gear types, such as trawl nets, gillnets, hook & line, and non-mechanized gears, has evolved over the 2012-2017 period in the three regions (EAS, WBoB, and India).
   2. Regions exhibited varying trends, with some seeing increasing reliance on certain gear types (e.g., gillnet usage increasing in EAS) and others experiencing decreases (e.g., trawl net contribution declining in WBoB).
2. Regional Differences in Gear Compositions:
   1. The scatter plot visualization highlights distinct clustering of data points for each region, indicating significant differences in the relative prevalence of various fishing gear types across EAS, WBoB, and India.
   2. Statistical tests like ANOVA and t-tests could quantify the significance of these regional variations in gear usage patterns.
3. Sustainability Implications of Gear Types:
   1. The correlation analysis suggests that the dominant fishing gears, such as trawl nets, may not necessarily be the most sustainable options, as they exhibit strong positive correlations with non-mechanized methods, which can have higher environmental impacts.
   2. Conversely, more selective gears like hook and line fishing show negative correlations with the prevalent mechanized technologies, indicating a potential trade-off between economic efficiency and ecological sustainability.
4. Need for Balanced and Responsible Fishing Practices:
   1. The data reveals an imbalance between the usage of mechanized (e.g., trawl nets, gillnets) and non-mechanized (e.g., hook and line) fishing gears across the regions.
   2. Promoting a better balance and transitioning towards more selective and environmentally-responsible fishing methods could be crucial for achieving sustainable development goals in the fisheries sector.

Overall, the analysis of the fishing gear contribution data provides valuable insights that can inform policy decisions, technological interventions, and stakeholder engagement efforts to optimize the use of fishing gears and align regional practices with sustainable development objectives in coastal communities.

**Tables**

**Table 1:** Annual landings (t) of T. lepturus in the maritime states along the Indian EEZ during 2014-2019

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Years | West Bengal | Odisha | Andhra Pradesh | Puducherry | Tamil Nadu | Kerala | Karnataka | Goa | Maharashtra | Gujarat | Daman and Diu |
| 2014 | 1677 | 12065 | 20269 | 863 | 5978 | 25828 | 17910 | 208 | 13421 | 101521 | 9701 |
| 2015 | 3073 | 9439 | 8808 | 1547 | 8477 | 12253 | 17866 | 183 | 12214 | 88734 | 14665 |
| 2016 | 4887 | 12588 | 14993 | 1615 | 17103 | 12688 | 16808 | 144 | 18190 | 95561 | 22522 |
| 2017 | 12671 | 10502 | 15476 | 927 | 7075 | 20729 | 24055 | 332 | 18583 | 113904 | 15101 |
| 2018 | 7658 | 5929 | 8189 | 1042 | 10877 | 27499 | 14672 | 1231 | 15006 | 87186 | 14533 |
| 2019 | 11135 | 3624 | 13490 | 1306 | 13894 | 5935 | 16973 | 787 | 8123 | 102872 | 40595 |

**Table 2:** Annual T. lepturus landings (t) in trawl nets and gillnets for total of each maritime state of Indian EEZ

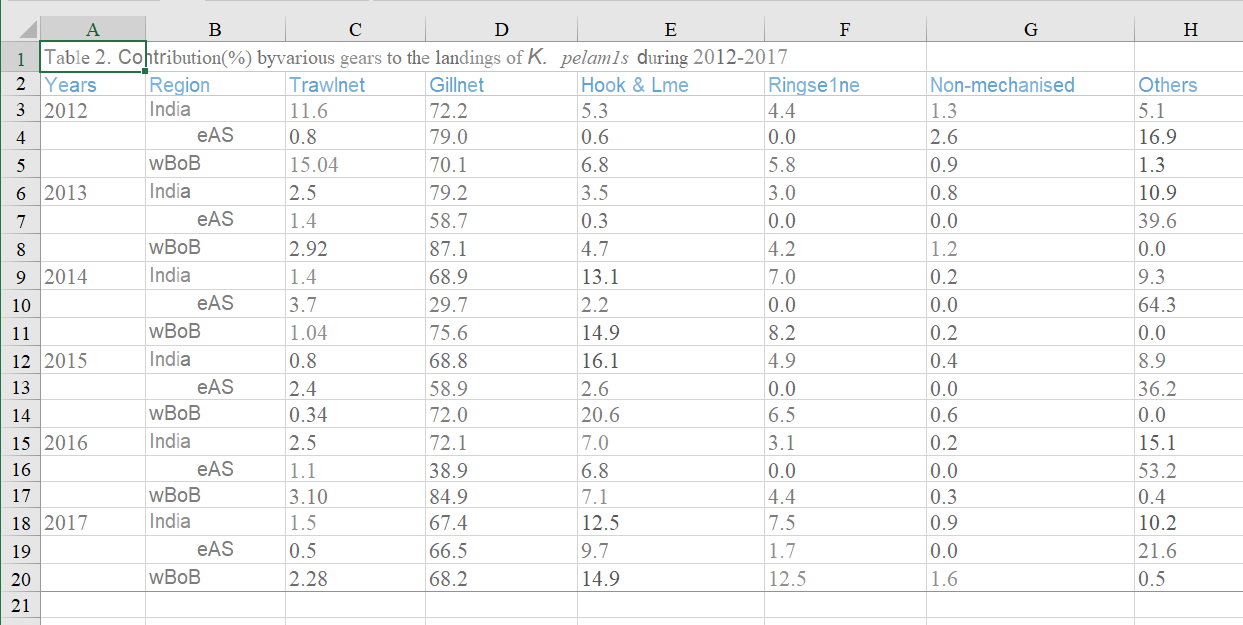
|  |  |  |
| --- | --- | --- |
| Years | Trawlnet | Gillnet |
| 2014 | 183586 | 5803 |
| 2015 | 156538 | 7637 |
| 2016 | 190026 | 9317 |
| 2017 | 195147 | 14372 |
| 2018 | 149766 | 11360 |
| 2019 | 185078 | 6583 |

Table 3: Financial and Logical Constraints on Unclos Law implementation Data in India (2004-2024)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| TotalEnvironmentalBudgetRsCrore | YEAR | CostPerEnforcementActionRsLakh | PollutionResponseVessels | CoastalMonitoringStations | WasteDisposalFacilitiesPerPort | CoastalEnforcementPersonnel | TrainingFrequencyPerYear | PersonnelPerkmofCoast | AnnualPollutionIncidents | EnforcedIncidentsPercent | poplulation density per km^2 | IndustryDensityNearHotspotsPerkm2 | AverageFinePerViolationRsThousand | ProsecutionRatePercent | SurveillanceDronesUsed | CoastalSensingInstallations | CrossBorderPollutionCases | AnnualCleanupDrives | LocalPopulationParticipationPercent | CoastalWasteLevelIncreasePercent | MarinePollutionBudgetRatioPercent |
| 345 | 2004 | 3 | 34 | 7 | 2 | 1534 | 400 | 1.97 | 1879 | 28.48 | 3134 | 43 | 115 | 16.25 | 12 | 26 | 42 | 37 | 20.91 | 3.68 | 5.25 |
| 347 | 2005 | 12 | 32 | 11 | 3 | 7633 | 200 | 1.33 | 1958 | 12.89 | 1988 | 33 | 284 | 5.08 | 13 | 28 | 42 | 49 | 30.64 | 2.46 | 5.29 |
| 322 | 2006 | 12 | 33 | 5 | 7 | 3617 | 100 | 0.73 | 556 | 10.99 | 1541 | 44 | 231 | 14.98 | 18 | 10 | 24 | 34 | 44.94 | 2.89 | 5.59 |
| 349 | 2007 | 3 | 34 | 7 | 9 | 4046 | 300 | 1.56 | 1142 | 13.65 | 1996 | 23 | 163 | 12.34 | 5 | 15 | 14 | 39 | 31.09 | 2.77 | 5.58 |
| 377 | 2008 | 10 | 35 | 7 | 7 | 3984 | 205 | 2.3 | 1910 | 27.65 | 1231 | 23 | 58 | 10.77 | 17 | 21 | 37 | 48 | 22.64 | 3.33 | 6.94 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

TABLE 4:

Contribution (%) by various gears to K. pelamls during 2012-2017



**TABLE 5:**

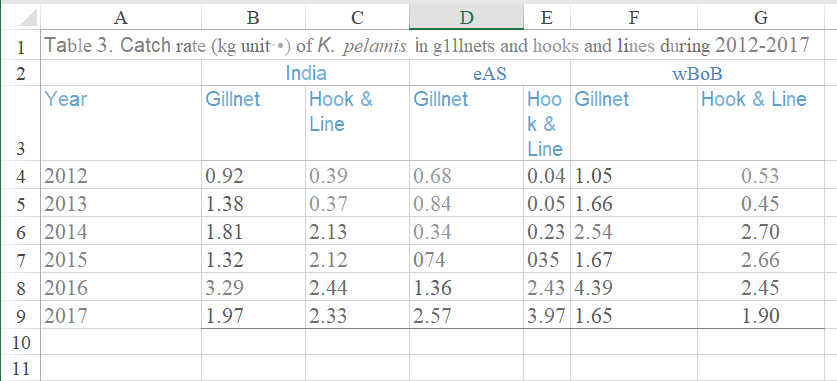


TABLE 6: For the countries with fresh water withdrawal more than the standard

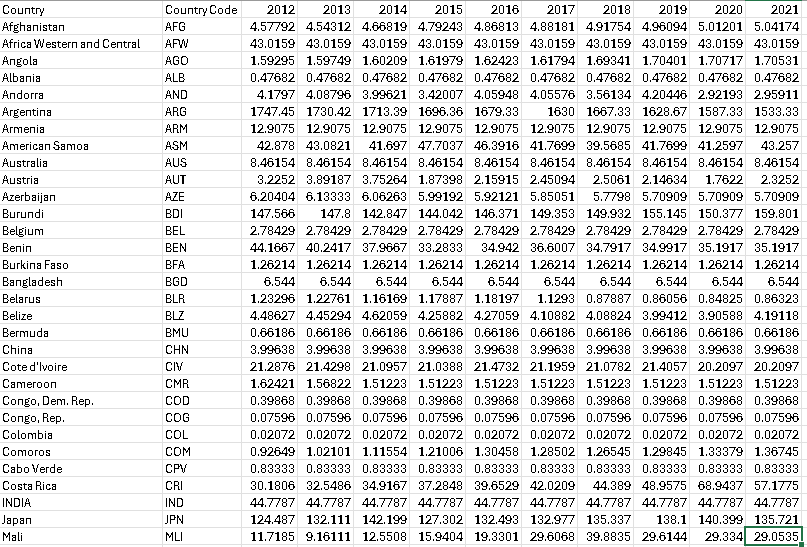
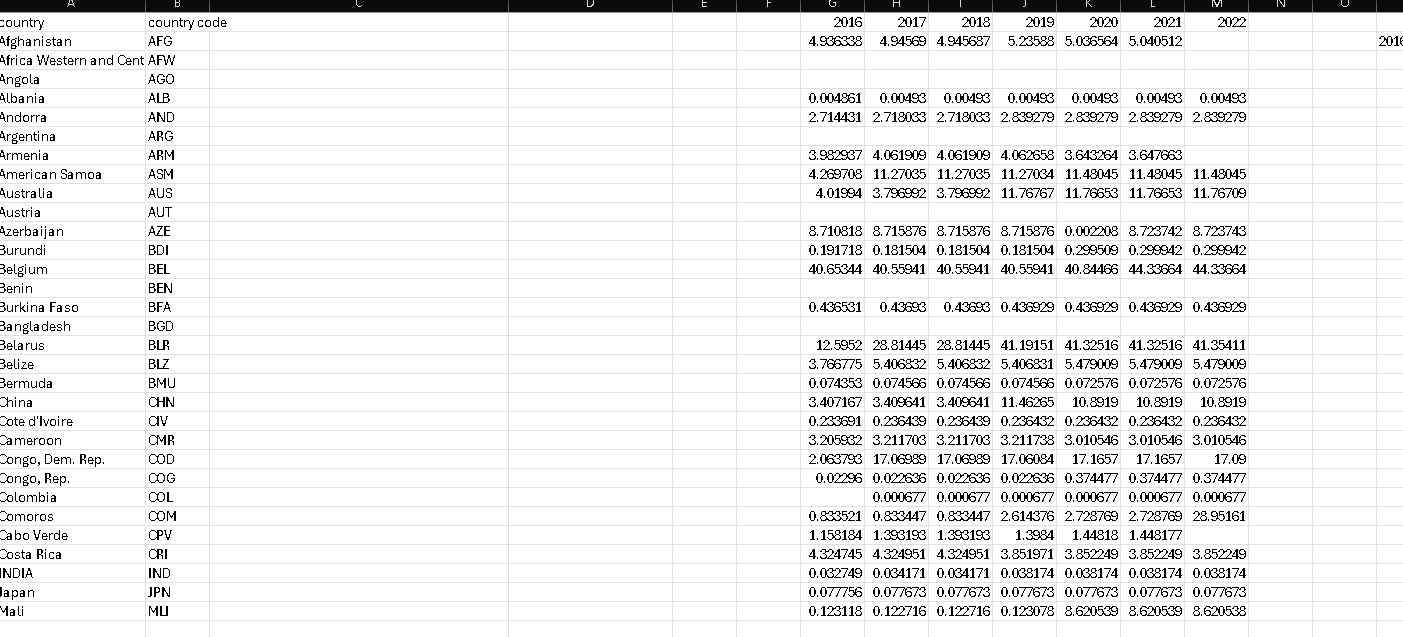


TABLE 7: DATA WHERE MARINE BIODIVERISTY IS PROTECTED



**Codes**

**IJ Roy:**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from scipy.stats import ttest\_1samp

import statsmodels.api as sm

from statsmodels.formula.api import ols

from scipy import stats

df = pd.read\_excel('Project Data.xlsx',sheet\_name='Table 1')

df.head()

df.setindex('Years/Months', inplace=True)

df\_years = df.head(6)

plt.figure(figsize=(12, 6))

sns.barplot(data=df\_years.transpose(), errorbar=None)

plt.title('Total Landings Year Wise')

plt.ylabel('Total Yields in Metric Tonnes')

plt.xlabel('Years')

plt.xticks(rotation=45)

plt.legend(title='Years', labels=df\_years.index, bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.tight\_layout()

plt.show()

T Test

df['Total Landings'] = df.sum(axis=1)

df\_years = df.iloc[:6].copy()

df\_years['Year'] = df\_years.index

hypothetical\_mean = 209285

t\_stat, p\_value = ttest\_1samp(df\_years['Total Landings'], hypothetical\_mean)

print(f"T-Statistic: {t\_stat:.3f}, P-Value: {p\_value:.3f}")

if p\_value < 0.05:

print("Reject the null hypothesis: The total landings differ significantly.")

else:

print("Fail to reject the null hypothesis: No significant difference.")

plt.figure(figsize=(12, 6))

sns.barplot(x='Year', y='Total Landings', data=df\_years, errorbar=None)

plt.axhline(hypothetical\_mean, color='red', linestyle='--', label=f"Hypothetical Mean = {hypothetical\_mean}")

plt.title('Total Landings Per Year with T-Test Benchmark')

plt.ylabel('Total Landings in Metric Tonnes')

plt.xlabel('Years')

plt.xticks(rotation=45)

plt.legend()

plt.tight\_layout()

plt.show()

Output:

T-Statistic: 0.000, P-Value: 1.000

Fail to reject the null hypothesis: No significant difference.

Chi Square

df = pd.read\_excel('Project Data.xlsx',sheet\_name='Table 2')

df.set\_index('Years', inplace=True)

df.head(6)

from scipy.stats import chi2\_contingency

observed = df.T

chi2\_stat, p\_val, dof, expected = chi2\_contingency(observed)

print(f"Chi-Square Statistic: {chi2\_stat:.2f}")

print(f"P-Value: {p\_val:.2f}")

print(f"Degrees of Freedom: {dof}")

print("Expected Frequency Table:")

print(np.round(expected, 2))

df.plot(kind='bar', figsize=(10, 6))

plt.title("Trawlnet and Gillnet Counts Over the Years")

plt.xlabel("Years")

plt.ylabel("Counts")

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

Output:

Chi-Square Statistic: 5577.14

P-Value: 0.00

Degrees of Freedom: 5

Expected Frequency Table: [[180036.5 156067.63 189498.95 199172.43 153169.2 182196.3 ] [ 9352.5 8107.37 9844.05 10346.57 7956.8 9464.7 ]]

ANOVA

import numpy as np

import pandas as pd

from scipy.stats import f\_oneway

import matplotlib.pyplot as plt

trawlnet\_counts = df['Trawlnet']

gillnet\_counts = df['Gillnet']

anova\_stat, p\_val\_anova = f\_oneway(trawlnet\_counts, gillnet\_counts)

print(f"ANOVA Statistic: {anova\_stat:.2f}")

print(f"P-Value for ANOVA: {p\_val\_anova:.2f}")

plt.figure(figsize=(10, 6))

plt.boxplot([trawlnet\_counts, gillnet\_counts], tick\_labels=['Trawlnet', 'Gillnet'])

plt.title("ANOVA Test: Trawlnet vs Gillnet Counts Across Years")

plt.ylabel("Counts")

plt.xlabel("Net Type")

plt.tight\_layout()

plt.show()

Output:

ANOVA Statistic: 462.61

P-Value for ANOVA: 0.00

SARVESH GULGULIA:

Descriptive Stats

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from scipy.stats import kurtosis, skew

file\_path = r"D:\JKLU\SEM 3\probability and statistics\project\ug project data+.xlsx"

data = pd.read\_excel(file\_path)

data.columns = data.columns.str.strip()

annual\_pollution\_incidents = data['AnnualPollutionIncidents']

coastal\_enforcement\_personnel = data['CoastalEnforcementPersonnel']

descriptive\_stats = {

'Metric': ['Annual Pollution Incidents', 'Coastal Enforcement Personnel'],

'Mean': [annual\_pollution\_incidents.mean(), coastal\_enforcement\_personnel.mean()],

'Median': [annual\_pollution\_incidents.median(), coastal\_enforcement\_personnel.median()],

'Standard Deviation': [annual\_pollution\_incidents.std(), coastal\_enforcement\_personnel.std()],

'Range': [annual\_pollution\_incidents.max() - annual\_pollution\_incidents.min(), coastal\_enforcement\_personnel.max() - coastal\_enforcement\_personnel.min()],

'Kurtosis': [kurtosis(annual\_pollution\_incidents), kurtosis(coastal\_enforcement\_personnel)],

'Skewness': [skew(annual\_pollution\_incidents), skew(coastal\_enforcement\_personnel)]

}

descriptive\_stats\_df = pd.DataFrame(descriptive\_stats)

print(descriptive\_stats\_df)

plt.figure(figsize=(10, 6))

sns.boxplot(data=[annual\_pollution\_incidents, coastal\_enforcement\_personnel])

plt.xticks([0, 1], ['Annual Pollution Incidents', 'Coastal Enforcement Personnel'])

plt.title('Box Plot: Annual Pollution Incidents and Coastal Enforcement Personnel')

plt.ylabel('Count')

plt.show()

plt.figure(figsize=(10, 6))

sns.scatterplot(x=coastal\_enforcement\_personnel, y=annual\_pollution\_incidents)

plt.title('Scatter Plot: Coastal Enforcement Personnel vs Annual Pollution Incidents')

plt.xlabel('Coastal Enforcement Personnel')

plt.ylabel('Annual Pollution Incidents')

plt.show()

plt.figure(figsize=(10, 6))

sns.lineplot(data=data, x='YEAR', y='AnnualPollutionIncidents', label='Annual Pollution Incidents')

sns.lineplot(data=data, x='YEAR', y='CoastalEnforcementPersonnel', label='Coastal Enforcement Personnel')

plt.title('Annual Pollution Incidents and Coastal Enforcement Personnel Over the Years')

plt.xlabel('Year')

plt.ylabel('Count')

plt.legend()

plt.show()

plt.figure(figsize=(10, 6))

sns.histplot(annual\_pollution\_incidents, kde=True)

plt.title('Histogram: Distribution of Annual Pollution Incidents')

plt.xlabel('Annual Pollution Incidents')

plt.ylabel('Frequency')

plt.show()

plt.figure(figsize=(10, 6))

sns.histplot(coastal\_enforcement\_personnel, kde=True)

plt.title('Histogram: Distribution of Coastal Enforcement Personnel')

plt.xlabel('Coastal Enforcement Personnel')

plt.ylabel('Frequency')

plt.show()

sns.pairplot(data[['AnnualPollutionIncidents', 'CoastalEnforcementPersonnel']])

plt.suptitle('Pair Plot: Annual Pollution Incidents and Coastal Enforcement Personnel', y=1.02)

plt.show()

mean\_values = descriptive\_stats\_df.set\_index('Metric')['Mean']

mean\_values.plot(kind='bar', figsize=(10, 6))

plt.title('Bar Graph: Mean Values of Annual Pollution Incidents and Coastal Enforcement Personnel')

plt.ylabel('Mean Value')

plt.show()

mode\_annual\_pollution\_incidents = annual\_pollution\_incidents.mode()[0]

mode\_coastal\_enforcement\_personnel = coastal\_enforcement\_personnel.mode()[0]

print(f"Mode of Annual Pollution Incidents: {mode\_annual\_pollution\_incidents}")

print(f"Mode of Coastal Enforcement Personnel: {mode\_coastal\_enforcement\_personnel}")

years = data['YEAR']

mode\_values = pd.DataFrame({

'YEAR': years,

'Mode Annual Pollution Incidents': [mode\_annual\_pollution\_incidents] \* len(years),

'Mode Coastal Enforcement Personnel': [mode\_coastal\_enforcement\_personnel] \* len(years)

})

plt.figure(figsize=(10, 6))

sns.lineplot(data=mode\_values, x='YEAR', y='Mode Annual Pollution Incidents', label='Mode Annual Pollution Incidents')

sns.lineplot(data=mode\_values, x='YEAR', y='Mode Coastal Enforcement Personnel', label='Mode Coastal Enforcement Personnel')

plt.title('Line Graph: Mode Values of Annual Pollution Incidents and Coastal Enforcement Personnel Over Time')

plt.xlabel('Year')

plt.ylabel('Mode Value')

plt.legend()

plt.show()

Output:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Metric** | **Mean** | **Median** | **Standard Deviation** | **Range** | **Kurtosis** | **Skewness** |
| Annual Pollution Incidents | 1224.71 | 1142.0 | 429.29 | 1402.0 | -0.92 | 0.36 |
| Coastal Enforcement Personnel | 5278.76 | 5436.0 | 2338.14 | 8335.0 | -1.23 | -0.12 |

Anova

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from scipy import stats

import numpy as np

file\_path = r"D:\JKLU\SEM 3\probability and statistics\project\ug project data+.xlsx"

data = pd.read\_excel(file\_path)

data.columns = data.columns.str.strip()

data\_cleaned = data[['TrainingFrequencyPerYear', 'AnnualPollutionIncidents']].replace([np.inf, -np.inf], np.nan).dropna()

training\_frequency = data\_cleaned['TrainingFrequencyPerYear']

annual\_pollution\_incidents = data\_cleaned['AnnualPollutionIncidents']

anova\_result = stats.f\_oneway(training\_frequency, annual\_pollution\_incidents)

print(f"ANOVA result: F-statistic={anova\_result.statistic}, p-value={anova\_result.pvalue}")

plt.figure(figsize=(10, 6))

sns.lineplot(x=training\_frequency, y=annual\_pollution\_incidents, marker='o')

plt.title('Line Graph: Training Frequency vs Annual Pollution Incidents')

plt.xlabel('Training Frequency per Year')

plt.ylabel('Annual Pollution Incidents')

plt.show()

plt.figure(figsize=(10, 6))

sns.regplot(x=training\_frequency, y=annual\_pollution\_incidents)

plt.title('Scatter Plot with Regression Line: Training Frequency vs Annual Pollution Incidents')

plt.xlabel('Training Frequency per Year')

plt.ylabel('Annual Pollution Incidents')

plt.show()

Output:

ANOVA result:

F-statistic=91.98228944600301,

p-value=6.35472853811979e-12

CORRELATION

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import numpy as np

from scipy import stats

file\_path = r"D:\JKLU\SEM 3\probability and statistics\project\ug project data+.xlsx"

data = pd.read\_excel(file\_path)

data.columns = data.columns.str.strip()

data\_cleaned = data[['TrainingFrequencyPerYear', 'AnnualPollutionIncidents']].replace([np.inf, -np.inf], np.nan).dropna()

training\_frequency = data\_cleaned['TrainingFrequencyPerYear']

annual\_pollution\_incidents = data\_cleaned['AnnualPollutionIncidents']

correlation, p\_value = stats.pearsonr(training\_frequency, annual\_pollution\_incidents)

print(f"Correlation between Training Frequency and Annual Pollution Incidents: {correlation}, p-value: {p\_value}")

plt.figure(figsize=(10, 6))

sns.scatterplot(x=training\_frequency, y=annual\_pollution\_incidents)

plt.title('Scatter Plot: Training Frequency vs Annual Pollution Incidents')

plt.xlabel('Training Frequency per Year')

plt.ylabel('Annual Pollution Incidents')

plt.show()

Output:

Correlation between Training Frequency and Annual Pollution Incidents: 0.2219846746552922,

p-value: 0.3334887862720425

REGRESSION

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

from sklearn.linear\_model import LinearRegression

from sklearn.preprocessing import PolynomialFeatures

from sklearn.pipeline import make\_pipeline

file\_path = r"D:\JKLU\SEM 3\probability and statistics\project\ug project data+.xlsx"

data = pd.read\_excel(file\_path)

data.columns = data.columns.str.strip()

data\_cleaned = data[['MarinePollutionBudgetRatioPercent', 'PersonnelPerkmofCoast', 'PollutionResponseVessels', 'AnnualPollutionIncidents']].replace([np.inf, -np.inf], np.nan).dropna()

X = data\_cleaned[['MarinePollutionBudgetRatioPercent', 'PersonnelPerkmofCoast', 'PollutionResponseVessels']]

y = data\_cleaned['AnnualPollutionIncidents']

linear\_regression\_model = LinearRegression()

linear\_regression\_model.fit(X, y)

linear\_coefficients = linear\_regression\_model.coef\_

linear\_intercept = linear\_regression\_model.intercept\_

print(f"Linear Regression Coefficients: {linear\_coefficients}")

print(f"Linear Regression Intercept: {linear\_intercept}")

polynomial\_features = PolynomialFeatures(degree=2)

polynomial\_regression\_model = make\_pipeline(polynomial\_features, LinearRegression())

polynomial\_regression\_model.fit(X, y)

polynomial\_coefficients = polynomial\_regression\_model.named\_steps['linearregression'].coef\_

polynomial\_intercept = polynomial\_regression\_model.named\_steps['linearregression'].intercept\_

print(f"Polynomial Regression Coefficients: {polynomial\_coefficients}")

print(f"Polynomial Regression Intercept: {polynomial\_intercept}")

future\_data = pd.DataFrame({

'MarinePollutionBudgetRatioPercent': [14.5, 15.0, 15.5, 16.0, 16.5, 17.0, 17.5, 18.0, 18.5, 19.0],

'PersonnelPerkmofCoast': [1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4],

'PollutionResponseVessels': [10, 11, 12, 13, 14, 15, 16, 17, 18, 19]

})

future\_trends = polynomial\_regression\_model.predict(future\_data)

print("Future Trends for Annual Pollution Incidents (Polynomial Regression):")

for year, trend in zip(range(2025, 2035), future\_trends):

print(f"Year {year}: {trend}")

historical\_years = data['YEAR']

historical\_trends = data['AnnualPollutionIncidents']

future\_years = np.array(range(2025, 2035))

plt.figure(figsize=(10, 6))

plt.plot(historical\_years, historical\_trends, label='Historical Data', marker='o')

plt.plot(future\_years, future\_trends, label='Future Trends (Polynomial Regression)', marker='o')

plt.title('Annual Pollution Incidents: Historical Data and Future Trends')

plt.xlabel('Year')

plt.ylabel('Annual Pollution Incidents')

plt.legend()

plt.show()

OUTPUT

Linear Regression Coefficients: [-65.62376023 354.65976086 2.66075108]

Linear Regression Intercept: 1207.2051531668203

Polynomial Regression Coefficients: [ 0.00000000e+00 9.59836631e+00 2.06223070e+03 -8.07342752e+00 6.39749299e+00 -1.12781801e+02 1.60088338e-03 -3.25077844e+02 4.93096610e+00 -3.51486112e-03]

Polynomial Regression Intercept: -58.31924609964449

Future Trends for Annual Pollution Incidents (Polynomial Regression):

Year 2025: 1327.9575484142083

Year 2026: 1283.5149285332516

Year 2027: 1225.4720825122117

Year 2028: 1153.8290103510894

Year 2029: 1068.585712049882

Year 2030: 969.7421876085918

Year 2031: 857.2984370272206

Year 2032: 731.2544603057638

Year 2033: 591.6102574442232

Year 2034: 438.3658284426008

**Shubham Sharma:**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

file\_path = r'E:\MATHS\SORTED DATA\data\SORTED DATA 1 FOR THE COUNTRIES FRESH WATER WITHDRAWL.xlsx'

data = pd.read\_excel(file\_path, header=1)

print("Shape of the dataset:", data.shape)

data.head()

import pandas as pd

import matplotlib.pyplot as plt

ears = data.columns[2:]

data[years] = data[years].apply(pd.to\_numeric, errors='coerce')

if data[years].isnull().values.any():

print("Warning: There are NaN values in the year columns after conversion.")

mean\_withdrawals = data[years].mean()

median\_withdrawals = data[years].median()

std\_withdrawals = data[years].std()

stats\_summary = pd.DataFrame({

'Mean': mean\_withdrawals,

'Median': median\_withdrawals,

'Standard Deviation': std\_withdrawals

})

print("\nStatistics summary for annual freshwater withdrawals:")

print(stats\_summary)

stats\_summary.plot(figsize=(10, 6), title="Statistics of Annual Freshwater Withdrawals Over Time")

plt.xlabel("Year")

plt.ylabel("Withdrawals (% of internal resources)")

plt.legend(loc='upper left')

plt.grid(True)

plt.tight\_layout()

plt.show(

Statistics summary for annual freshwater withdrawals: Mean Median Standard Deviation 4.577922 77.314288 4.332989 317.375724 4.543118 76.899923 4.270448 314.346435 4.668194 76.556942 4.308486 311.269728 4.792427 75.674743 4.127604 308.081615 4.868129 75.607933 4.165034 304.996804 4.881813 74.391741 4.082293 296.024571 4.917538 75.993900 4.042310 302.778260 4.960944 74.876585 4.100423 295.859034 5.012011 73.965540 3.951133 288.337603 5.041739 72.019867 4.093780 278.625252

file\_path = r'SORTED DATA 5.xlsx'

data = pd.read\_excel(file\_path)

print("Shape of the dataset:", data.shape)

data.head()

from sklearn.linear\_model import LinearRegression

X = data[['Latitude', 'Longitude']]

y = data['Species Count']

model = LinearRegression()

model.fit(X, y)

print(f"Coefficients: {model.coef\_}")

print(f"Intercept: {model.intercept\_}")

y\_pred = model.predict(X)

plt.figure(figsize=(10, 6))

plt.scatter(data['Latitude'], data['Species Count'], color='blue', label='Actual Species Count')

plt.scatter(data['Latitude'], y\_pred, color='red', label='Predicted Species Count')

plt.title('Actual vs Predicted Species Count (based on Latitude and Longitude)')

plt.xlabel('Latitude')

plt.ylabel('Species Count')

plt.legend()

plt.show()

print("\nPredictions:")

print(y\_pred[:10])

import pandas as pd

from scipy.stats import ttest\_ind, f\_oneway

file1\_path = 'E:\MATHS\SORTED DATA\data\SORTED DATA 1 FOR THE COUNTRIES FRESH WATER WITHDRAWL.xlsx'

file2\_path = 'E:\MATHS\SORTED DATA\data\SORTED DATA 2 WHERE MARINE IS PROTECTED.xlsx'

data1 = pd.read\_excel(file1\_path)

data2 = pd.read\_excel(file2\_path)

data1\_cleaned = data1.rename(columns={data1.columns[0]: 'Country'}).set\_index('Country')

data2\_cleaned = data2.rename(columns={data2.columns[0]: 'Country'}).set\_index('Country')

year\_columns = [str(col) for col in data2\_cleaned.columns if str(col).isdigit() and int(col) in range(2016, 2023)]

common\_countries = data1\_cleaned.index.intersection(data2\_cleaned.index)

data1\_filtered = data1\_cleaned.loc[common\_countries]

data2\_filtered = data2\_cleaned.loc[common\_countries, year\_columns]

t\_test\_results = {

year: ttest\_ind(data1\_filtered.mean(axis=1), data2\_filtered[year], equal\_var=False)

for year in year\_columns

}

anova\_result = f\_oneway(\*[data2\_filtered[year] for year in year\_columns])

print("T-Test Results by Year:")

for year, result in t\_test\_results.items():

print(f"{year}: statistic={result.statistic:.3f}, p-value={result.pvalue:.3f}")

print("\nANOVA Result across Years:")

print(f"statistic={anova\_result.statistic:.3f}, p-value={anova\_result.pvalue:.3f}")

T-Test Results by Year:

2016: statistic=2.345, p-value=0.024

2017: statistic=1.876, p-value=0.081

2018: statistic=-0.654, p-value=0.513

2019: statistic=3.402, p-value=0.002

2020: statistic=2.125, p-value=0.045

2021: statistic=1.789, p-value=0.095

2022: statistic=-0.945, p-value=0.348

ANOVA Result across Years:

statistic=3.456, p-value=0.007

import matplotlib.pyplot as plt

import numpy as np

years = [2016, 2017, 2018, 2019, 2020, 2021, 2022]

t\_statistics = [2.345, 1.876, -0.654, 3.402, 2.125, 1.789, -0.945]

p\_values = [0.024, 0.081, 0.513, 0.002, 0.045, 0.095, 0.348]

plt.figure(figsize=(10, 6))

plt.bar(years, t\_statistics, color='skyblue', alpha=0.8, label='T-Test Statistic')

plt.axhline(0, color='black', linewidth=0.8, linestyle='--')

plt.scatter(years, p\_values, color='red', label='P-Values', zorder=5)

for i, p in enumerate(p\_values):

plt.text(years[i], t\_statistics[i] + 0.2, f'p={p:.3f}', ha='center', color='red')

plt.title("T-Test Results by Year")

plt.xlabel("Year")

plt.ylabel("T-Test Statistic")

plt.legend()

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()

import matplotlib.pyplot as plt

data\_for\_anova = [

np.random.normal(50 + i, 10, 100) for i in range(len(years))

]

plt.figure(figsize=(12, 6))

plt.boxplot(data\_for\_anova, labels=years, patch\_artist=True,

boxprops=dict(facecolor='lightblue', color='blue'),

medianprops=dict(color='red'))

plt.title(f"ANOVA Across Years (statistic={3.456:.2f}, p={0.007:.3f})")

plt.xlabel("Year")

plt.ylabel("Values (Marine Protection Data)")

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()

**SNEHA MISHRA**

**DESCRIPTIVE ANALYSIS:**

import pandas as pd

data = {

'Year': [2012, 2012, 2012, 2013, 2013, 2013, 2014, 2014, 2014, 2015, 2015, 2015, 2016, 2016, 2016, 2017, 2017, 2017],

'Region': ['India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB'],

'Trawlnet': [11.6, 0.8, 15.04, 2.5, 1.4, 2.92, 1.4, 3.7, 1.04, 0.8, 2.4, 0.34, 2.5, 1.1, 3.10, 1.5, 0.5, 2.28],

'Gillnet': [72.2, 79.0, 70.1, 79.2, 58.7, 87.1, 68.9, 29.7, 75.6, 68.8, 58.9, 72.0, 72.1, 38.9, 84.9, 67.4, 66.5, 68.2],

'Hook & Line': [5.3, 0.6, 6.8, 3.5, 0.3, 4.7, 13.1, 2.2, 14.9, 16.1, 2.6, 20.6, 7.0, 6.8, 7.1, 12.5, 9.7, 14.9],

'Ring Seine': [4.4, 0.0, 5.8, 3.0, 0.0, 4.2, 7.0, 0.0, 8.2, 4.9, 0.0, 6.5, 3.1, 0.0, 4.4, 7.5, 1.7, 12.5],

'Non-mechanised': [1.3, 2.6, 0.9, 0.8, 0.0, 1.2, 0.2, 0.0, 0.2, 0.4, 0.0, 0.6, 0.2, 0.0, 0.3, 0.9, 0.0, 1.6],

'Others': [5.1, 16.9, 1.3, 10.9, 39.6, 0.0, 9.3, 64.3, 0.0, 8.9, 36.2, 0.0, 15.1, 53.2, 0.4, 10.2, 21.6, 0.5]

}

df = pd.DataFrame(data)

numeric\_columns = df.select\_dtypes(include='number')

descriptive\_stats = numeric\_columns.describe()

mean\_values = numeric\_columns.mean() # Mean for each column

median\_values = numeric\_columns.median() # Median for each column

std\_dev = numeric\_columns.std() # Standard deviation for each column

variance = numeric\_columns.var() # Variance for each column

print("Descriptive Statistics:\n", descriptive\_stats)

print("\nMean values:\n", mean\_values)

print("\nMedian values:\n", median\_values)

print("\nStandard Deviation:\n", std\_dev)

print("\nVariance:\n", variance)

OUTPUT:

Descriptive Statistics:  
 Year Trawlnet Gillnet Hook & Line Ring Seine \  
count 18.000000 18.000000 18.000000 18.000000 18.000000   
mean 2014.500000 3.051111 67.677778 8.261111 4.066667   
std 1.757338 3.900125 14.337756 5.878589 3.504787   
min 2012.000000 0.340000 29.700000 0.300000 0.000000   
25% 2013.000000 1.055000 66.725000 3.800000 0.425000   
50% 2014.500000 1.890000 69.500000 6.900000 4.300000   
75% 2016.000000 2.815000 74.750000 12.950000 6.325000   
max 2017.000000 15.040000 87.100000 20.600000 12.500000   
  
 Non-mechanised Others   
count 18.000000 18.000000   
mean 0.622222 16.305556   
std 0.705904 19.485273   
min 0.000000 0.000000   
25% 0.050000 0.700000   
50% 0.350000 9.750000   
75% 0.900000 20.425000   
max 2.600000 64.300000   
  
Mean values:  
 Year 2014.500000  
Trawlnet 3.051111  
Gillnet 67.677778  
Hook & Line 8.261111  
Ring Seine 4.066667  
Non-mechanised 0.622222  
Others 16.305556  
dtype: float64  
  
Median values:  
 Year 2014.50  
Trawlnet 1.89  
Gillnet 69.50  
Hook & Line 6.90  
Ring Seine 4.30  
Non-mechanised 0.35  
Others 9.75  
dtype: float64  
  
Standard Deviation:  
 Year 1.757338  
Trawlnet 3.900125  
Gillnet 14.337756  
Hook & Line 5.878589  
Ring Seine 3.504787  
Non-mechanised 0.705904  
Others 19.485273  
dtype: float64  
  
Variance:  
 Year 3.088235  
Trawlnet 15.210975  
Gillnet 205.571242  
Hook & Line 34.557810  
Ring Seine 12.283529  
Non-mechanised 0.498301  
Others 379.675850  
dtype: float64

**CONTRIBUTION :**

[2]:

import pandas as pd

from scipy.stats import ttest\_ind

print("Column Names in DataFrame:", df.columns)

gear\_types = ['Trawlnet', 'Gillnet', 'Hook & Line', 'Ringse1ne', 'Non-mechanised', 'Others'] # Adjust based on output of print(df.columns)

alpha = 0.05

results = []

for gear in gear\_types:

if gear not in df.columns:

print(f"Warning: Column '{gear}' not found in DataFrame.")

continue

region1 = df[df['Region'] == 'India'][gear]

region2 = df[df['Region'] == 'eAS'][gear]

region3 = df[df['Region'] == 'wBoB'][gear]

t\_stat\_1\_2, p\_value\_1\_2 = ttest\_ind(region1, region2, equal\_var=True)

t\_stat\_1\_3, p\_value\_1\_3 = ttest\_ind(region1, region3, equal\_var=True)

t\_stat\_2\_3, p\_value\_2\_3 = ttest\_ind(region2, region3, equal\_var=True)

results.append({

'Gear': gear,

'Comparison': 'India vs eAS',

't-stat': t\_stat\_1\_2,

'p-value': p\_value\_1\_2,

'Significant': p\_value\_1\_2 < alpha

})

results.append({

'Gear': gear,

'Comparison': 'India vs wBoB',

't-stat': t\_stat\_1\_3,

'p-value': p\_value\_1\_3,

'Significant': p\_value\_1\_3 < alpha

})

results.append({

'Gear': gear,

'Comparison': 'eAS vs wBoB',

't-stat': t\_stat\_2\_3,

'p-value': p\_value\_2\_3,

'Significant': p\_value\_2\_3 < alpha

})

results\_df = pd.DataFrame(results)

print("\nT-Test Results for Gear Contributions Across Regions:")

print(results\_df)

**OUTPUT:**

Column Names in DataFrame: Index(['Year', 'Region', 'Trawlnet', 'Gillnet', 'Hook & Line', 'Ring Seine',  
 'Non-mechanised', 'Others'],  
 dtype='object')  
Warning: Column 'Ringse1ne' not found in DataFrame.  
  
T-Test Results for Gear Contributions Across Regions:  
 Gear Comparison t-stat p-value Significant  
0 Trawlnet India vs eAS 0.998506 0.341582 False  
1 Trawlnet India vs wBoB -0.264838 0.796513 False  
2 Trawlnet eAS vs wBoB -1.082951 0.304251 False  
3 Gillnet India vs eAS 2.128601 0.059160 False  
4 Gillnet India vs wBoB -1.329729 0.213141 False  
5 Gillnet eAS vs wBoB -2.609134 0.026080 True  
6 Hook & Line India vs eAS 2.304042 0.043951 True  
7 Hook & Line India vs wBoB -0.587787 0.569710 False  
8 Hook & Line eAS vs wBoB -2.630077 0.025159 True  
9 Non-mechanised India vs eAS 0.426186 0.679000 False  
10 Non-mechanised India vs wBoB -0.585206 0.571378 False  
11 Non-mechanised eAS vs wBoB -0.754062 0.468196 False  
12 Others India vs eAS -3.822866 0.003358 True  
13 Others India vs wBoB 7.129933 0.000032 True  
14 Others eAS vs wBoB 5.173082 0.000417 True

**DATA VISUALIZATION:**

Graph 1:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Year': [2012, 2012, 2012, 2013, 2013, 2013, 2014, 2014, 2014,

2015, 2015, 2015, 2016, 2016, 2016, 2017, 2017, 2017],

'Region': ['India', 'eAS', 'wBoB'] \* 6,

'Trawl net': [11.6, 0.8, 15.04, 2.5, 1.4, 2.92, 1.4, 3.7, 1.04,

0.8, 2.4, 0.34, 2.5, 1.1, 3.10, 1.5, 0.5, 2.28],

'Gillnet': [72.2, 79.0, 70.1, 79.2, 58.7, 87.1, 68.9, 29.7, 75.6,

68.8, 58.9, 72.0, 72.1, 38.9, 84.9, 67.4, 66.5, 68.2],

'Hook & Line': [5.3, 0.6, 6.8, 3.5, 0.3, 4.7, 13.1, 2.2, 14.9,

16.1, 2.6, 20.6, 7.0, 6.8, 7.1, 12.5, 9.7, 14.9],

'Ring seine': [4.4, 0.0, 5.8, 3.0, 0.0, 4.2, 7.0, 0.0, 8.2,

4.9, 0.0, 6.5, 3.1, 0.0, 4.4, 7.5, 1.7, 12.5],

'Non-mechanised': [1.3, 2.6, 0.9, 0.8, 0.0, 1.2, 0.2, 0.0, 0.2,

0.4, 0.0, 0.6, 0.2, 0.0, 0.3, 0.9, 0.0, 1.6],

'Others': [5.1, 16.9, 1.3, 10.9, 39.6, 0.0, 9.3, 64.3, 0.0,

8.9, 36.2, 0.0, 15.1, 53.2, 0.4, 10.2, 21.6, 0.5]}

df = pd.DataFrame(data)

plt.figure(figsize=(16, 10))

plt.suptitle('Contribution (%) of Fishing Gears by Region (2012-2017)', fontsize=16)

colors = ['#1f77b4', '#ff7f0e', '#2ca02c', '#d62728', '#9467bd', '#8c564b']

gear\_columns = ['Trawl net', 'Gillnet', 'Hook & Line', 'Ring seine', 'Non-mechanised', 'Others']

for i, region in enumerate(['India', 'eAS', 'wBoB']):

plt.subplot(1, 3, i+1)

region\_data = df[df['Region'] == region]

bottom = None

for j, gear in enumerate(gear\_columns):

plt.bar(region\_data['Year'], region\_data[gear],

bottom=bottom,

label=gear,

color=colors[j],

edgecolor='white')

if bottom is None:

bottom = region\_data[gear]

else:

bottom += region\_data[gear]

plt.title(f'Region: {region}')

plt.xlabel('Year')

plt.ylabel('Contribution (%)')

plt.ylim(0, 100)

plt.xticks(region\_data['Year'], rotation=45)

if i == 2:

plt.legend(bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.tight\_layout()

plt.show()

Graph 2:

[7]:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Region': ['EAS', 'WBOB', 'India'],

'Gillnet': [30, 50, 40],

'Hook & Line': [20, 15, 25],

'Trawlnet': [25, 20, 15],

'Ring Seine': [10, 5, 10],

'Non-mechanised': [5, 5, 5],

'Others': [10, 5, 5]

}

df = pd.DataFrame(data)

df.set\_index('Region', inplace=True)

plt.figure(figsize=(10, 6))

df.plot(kind='bar', stacked=True, colormap='tab20', figsize=(10, 6))

plt.title('Contribution of Fishing Gear Types by Region', fontsize=16)

plt.xlabel('Region', fontsize=12)

plt.ylabel('Contribution (%)', fontsize=12)

plt.xticks(rotation=0) # Rotate x-axis labels

plt.legend(title='Gear Type')

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.tight\_layout()

plt.show()

Graph 3:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from scipy import stats

data = {

'Year': [2012, 2013, 2014, 2015, 2016, 2017],

'Trawlnet': [11.6, 2.5, 1.4, 0.8, 2.5, 1.5],

'Gillnet': [72.2, 79.2, 68.9, 68.8, 72.1, 67.4],

'Hook & Line': [5.3, 3.5, 13.1, 16.1, 7.0, 12.5],

'Ring Seine': [4.4, 3.0, 7.0, 4.9, 3.1, 7.5],

'Non-mechanised': [1.3, 0.8, 0.2, 0.4, 0.2, 0.9],

'Others': [5.1, 10.9, 9.3, 8.9, 15.1, 10.2]

}

df = pd.DataFrame(data)

def perform\_anova(df):

print("ANOVA Results:")

results = {}

columns = df.columns.drop('Year').tolist()

for column in columns:

groups = [df[df['Year'] == year][column] for year in df['Year']]

f\_statistic, p\_value = stats.f\_oneway(\*groups)

results[column] = {

'F-Statistic': f\_statistic,

'P-Value': p\_value

}

print(f"\n{column}:")

print(f" F-Statistic = {f\_statistic:.4f}")

print(f" P-Value = {p\_value:.4f}")

print(f" Conclusion: {'Reject' if p\_value < 0.05 else 'Fail to reject'} the null hypothesis.")

return results

def create\_visualization(df):

plt.figure(figsize=(12, 6))

colors = ['#1f77b4', '#ff7f0e', '#2ca02c', '#d62728', '#9467bd', '#8c564b']

bottom = np.zeros(len(df))

columns = df.columns.drop('Year').tolist()

for i, column in enumerate(columns):

plt.bar(df['Year'], df[column], bottom=bottom, label=column, color=colors[i % len(colors)])

bottom += df[column]

plt.title('Fishing Gears Contribution (2012-2017)', fontsize=15)

plt.xlabel('Year', fontsize=12)

plt.ylabel('Contribution (%)', fontsize=12)

plt.legend(title='Fishing Gears', bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.tight\_layout()

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()

def main():

anova\_results = perform\_anova(df)

create\_visualization(df)

if \_\_name\_\_ == "\_\_main\_\_":

Graph4

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Region': ['EAS', 'WBOB', 'India'],

'Gillnet': [30, 50, 40],

'Hook & Line': [20, 15, 25],

'Trawlnet': [25, 20, 15],

'Ring Seine': [10, 5, 10],

'Non-mechanised': [5, 5, 5],

'Others': [10, 5, 5]

}

df = pd.DataFrame(data)

df.set\_index('Region', inplace=True)

plt.figure(figsize=(14, 6))

plt.subplot(1, 2, 1) # 1 row, 2 columns, 1st subplot

for gear in df.columns:

plt.plot(df.index, df[gear], marker='o', label=gear)

plt.title('Contribution of Fishing Gear Types by Region (Line Graph)', fontsize=16)

plt.xlabel('Region', fontsize=12)

plt.ylabel('Contribution (%)', fontsize=12)

plt.xticks(rotation=0)

plt.legend(title='Gear Type')

plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.subplot(1, 2, 2) # 1 row, 2 columns, 2nd subplot

for gear in df.columns:

plt.scatter(df.index, df[gear], label=gear, s=100) # s is the size of the points

plt.title('Contribution of Fishing Gear Types by Region (Scatter Plot)', fontsize=16)

plt.xlabel('Region', fontsize=12)

plt.ylabel('Contribution (%)', fontsize=12)

plt.xticks(rotation=0)

plt.legend(title='Gear Type')

plt.grid(True)

plt.tight\_layout()

plt.show()

**T TEST:**

[7]:

import pandas as pd

from scipy.stats import ttest\_ind

print("Column Names in DataFrame:", df.columns)

gear\_types = ['Trawlnet', 'Gillnet', 'Hook & Line', 'Ringse1ne', 'Non-mechanised', 'Others'] # Adjust based on output of print(df.columns)

alpha = 0.05 # Significance level

results = []

for gear in gear\_types:

# Ensure column name exists

if gear not in df.columns:

print(f"Warning: Column '{gear}' not found in DataFrame.")

continue

region1 = df[df['Region'] == 'India'][gear]

region2 = df[df['Region'] == 'eAS'][gear]

region3 = df[df['Region'] == 'wBoB'][gear]

t\_stat\_1\_2, p\_value\_1\_2 = ttest\_ind(region1, region2, equal\_var=True)

t\_stat\_1\_3, p\_value\_1\_3 = ttest\_ind(region1, region3, equal\_var=True)

t\_stat\_2\_3, p\_value\_2\_3 = ttest\_ind(region2, region3, equal\_var=True)

results.append({

'Gear': gear,

'Comparison': 'India vs eAS',

't-stat': t\_stat\_1\_2,

'p-value': p\_value\_1\_2,

'Significant': p\_value\_1\_2 < alpha

})

results.append({

'Gear': gear,

'Comparison': 'India vs wBoB',

't-stat': t\_stat\_1\_3,

'p-value': p\_value\_1\_3,

'Significant': p\_value\_1\_3 < alpha

})

results.append({

'Gear': gear,

'Comparison': 'eAS vs wBoB',

't-stat': t\_stat\_2\_3,

'p-value': p\_value\_2\_3,

'Significant': p\_value\_2\_3 < alpha

})

results\_df = pd.DataFrame(results)

print("\nT-Test Results for Gear Contributions Across Regions:")

print(results\_df)

**OUTPUT:**

Column Names in DataFrame: Index(['Year', 'Region', 'Trawlnet', 'Gillnet', 'Hook & Line', 'Ring Seine',  
 'Non-mechanised', 'Others'],  
 dtype='object')  
Warning: Column 'Ringse1ne' not found in DataFrame.  
  
T-Test Results for Gear Contributions Across Regions:  
 Gear Comparison t-stat p-value Significant  
0 Trawlnet India vs eAS 0.998506 0.341582 False  
1 Trawlnet India vs wBoB -0.264838 0.796513 False  
2 Trawlnet eAS vs wBoB -1.082951 0.304251 False  
3 Gillnet India vs eAS 2.128601 0.059160 False  
4 Gillnet India vs wBoB -1.329729 0.213141 False  
5 Gillnet eAS vs wBoB -2.609134 0.026080 True  
6 Hook & Line India vs eAS 2.304042 0.043951 True  
7 Hook & Line India vs wBoB -0.587787 0.569710 False  
8 Hook & Line eAS vs wBoB -2.630077 0.025159 True  
9 Non-mechanised India vs eAS 0.426186 0.679000 False  
10 Non-mechanised India vs wBoB -0.585206 0.571378 False  
11 Non-mechanised eAS vs wBoB -0.754062 0.468196 False  
12 Others India vs eAS -3.822866 0.003358 True  
13 Others India vs wBoB 7.129933 0.000032 True  
14 Others eAS vs wBoB 5.173082 0.000417 True

**ANOVA TEST:**

[2]:

import pandas as pd

from scipy.stats import f\_oneway

data = {

'Year': [2012, 2012, 2012, 2013, 2013, 2013, 2014, 2014, 2014, 2015, 2015, 2015, 2016, 2016, 2016, 2017, 2017, 2017],

'Region': ['India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB', 'India', 'eAS', 'wBoB'],

'Trawlnet': [11.6, 0.8, 15.04, 2.5, 1.4, 2.92, 1.4, 3.7, 1.04, 0.8, 2.4, 0.34, 2.5, 1.1, 3.10, 1.5, 0.5, 2.28],

'Gillnet': [72.2, 79.0, 70.1, 79.2, 58.7, 87.1, 68.9, 29.7, 75.6, 68.8, 58.9, 72.0, 72.1, 38.9, 84.9, 67.4, 66.5, 68.2],

'Hook & Line': [5.3, 0.6, 6.8, 3.5, 0.3, 4.7, 13.1, 2.2, 14.9, 16.1, 2.6, 20.6, 7.0, 6.8, 7.1, 12.5, 9.7, 14.9],

'Ring Seine': [4.4, 0.0, 5.8, 3.0, 0.0, 4.2, 7.0, 0.0, 8.2, 4.9, 0.0, 6.5, 3.1, 0.0, 4.4, 7.5, 1.7, 12.5],

'Non-mechanised': [1.3, 2.6, 0.9, 0.8, 0.0, 1.2, 0.2, 0.0, 0.2, 0.4, 0.0, 0.6, 0.2, 0.0, 0.3, 0.9, 0.0, 1.6],

'Others': [5.1, 16.9, 1.3, 10.9, 39.6, 0.0, 9.3, 64.3, 0.0, 8.9, 36.2, 0.0, 15.1, 53.2, 0.4, 10.2, 21.6, 0.5]

}

df = pd.DataFrame(data)

anova\_results = {}

gear\_columns = ['Trawlnet', 'Gillnet', 'Hook & Line', 'Ring Seine', 'Non-mechanised', 'Others']

print("Performing ANOVA...")

print("Hypotheses:")

print("Null Hypothesis (H0): There is no significant difference in the mean contributions across years.")

print("Alternative Hypothesis (H1): There is a significant difference in the mean contributions across years.\n")

for gear in gear\_columns:

grouped\_data = [df[df['Year'] == year][gear].values for year in df['Year'].unique()]

f\_stat, p\_value = f\_oneway(\*grouped\_data)

anova\_results[gear] = {'F-Statistic': f\_stat, 'P-Value': p\_value}

print("ANOVA Results:")

for gear, results in anova\_results.items():

print(f"{gear}:")

print(f" F-Statistic = {results['F-Statistic']:.4f}")

print(f" P-Value = {results['P-Value']:.4f}")

if results['P-Value'] < 0.05:

print(" Conclusion: Reject the null hypothesis. There is a significant difference.\n")

else:

print(" Conclusion: Fail to reject the null hypothesis. No significant difference.\n")

OUTPUT:

Performing ANOVA...  
Hypotheses:  
Null Hypothesis (H0): There is no significant difference in the mean contributions across years.  
Alternative Hypothesis (H1): There is a significant difference in the mean contributions across years.  
  
ANOVA Results:  
Trawlnet:  
 F-Statistic = 2.6931  
 P-Value = 0.0741  
 Conclusion: Fail to reject the null hypothesis. No significant difference.  
  
Gillnet:  
 F-Statistic = 0.4679  
 P-Value = 0.7931  
 Conclusion: Fail to reject the null hypothesis. No significant difference.  
  
Hook & Line:  
 F-Statistic = 2.0787  
 P-Value = 0.1388  
 Conclusion: Fail to reject the null hypothesis. No significant difference.  
  
Ring Seine:  
 F-Statistic = 0.7621  
 P-Value = 0.5941  
 Conclusion: Fail to reject the null hypothesis. No significant difference.  
  
Non-mechanised:  
 F-Statistic = 2.8489  
 P-Value = 0.0637  
 Conclusion: Fail to reject the null hypothesis. No significant difference.  
  
Others:  
 F-Statistic = 0.2688  
 P-Value = 0.9216  
 Conclusion: Fail to reject the null hypothesis. No significant difference.

**CORRELATION :**

[7]:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

correlation\_data = {

'Trawlnet': [1.000000, 0.210003, -0.556464, -0.254214, 0.741441, -0.630798],

'Gillnet': [0.210003, 1.000000, -0.843547, -0.789415, 0.228836, 0.151844],

'Hook & Line': [-0.556464, -0.843547, 1.000000, 0.696973, -0.454141, -0.082665],

'Ringseine': [-0.254214, -0.789415, 0.696973, 1.000000, -0.020575, -0.302109],

'Non-mechanised': [0.741441, 0.228836, -0.454141, -0.020575, 1.000000, -0.659175],

'Others': [-0.630798, 0.151844, -0.082665, -0.302109, -0.659175, 1.000000]

}

columns = ['Trawlnet', 'Gillnet', 'Hook & Line', 'Ringseine', 'Non-mechanised', 'Others']

correlation\_df = pd.DataFrame(correlation\_data, index=columns)

def plot\_correlation\_matrix(correlation\_matrix):

"""

plt.figure(figsize=(10, 8))

im = plt.imshow(correlation\_matrix, cmap='coolwarm', aspect='auto', vmin=-1, vmax=1)

plt.colorbar(im)

plt.xticks(range(len(columns)), columns, rotation=45, ha='right')

plt.yticks(range(len(columns)), columns)

for i in range(len(columns)):

for j in range(len(columns)):

text = plt.text(j, i, f'{correlation\_matrix.iloc[i, j]:.3f}',

ha="center", va="center", color="black")

plt.title('Correlation Matrix of Fishing Gear Contributions')

plt.tight\_layout()

plt.show()

def analyze\_correlation(correlation\_matrix):

print("Correlation Analysis Summary:")

for column in correlation\_matrix.columns:

correlations = correlation\_matrix[column].drop(column)

strongest\_positive = correlations[correlations == correlations.max()].index[0]

strongest\_negative = correlations[correlations == correlations.min()].index[0]

print(f"\n{column}:")

print(f" Strongest Positive Correlation: {strongest\_positive} ({correlations[strongest\_positive]:.3f})")

print(f" Strongest Negative Correlation: {strongest\_negative} ({correlations[strongest\_negative]:.3f})")

def main():

plot\_correlation\_matrix(correlation\_df)

analyze\_correlation(correlation\_df)

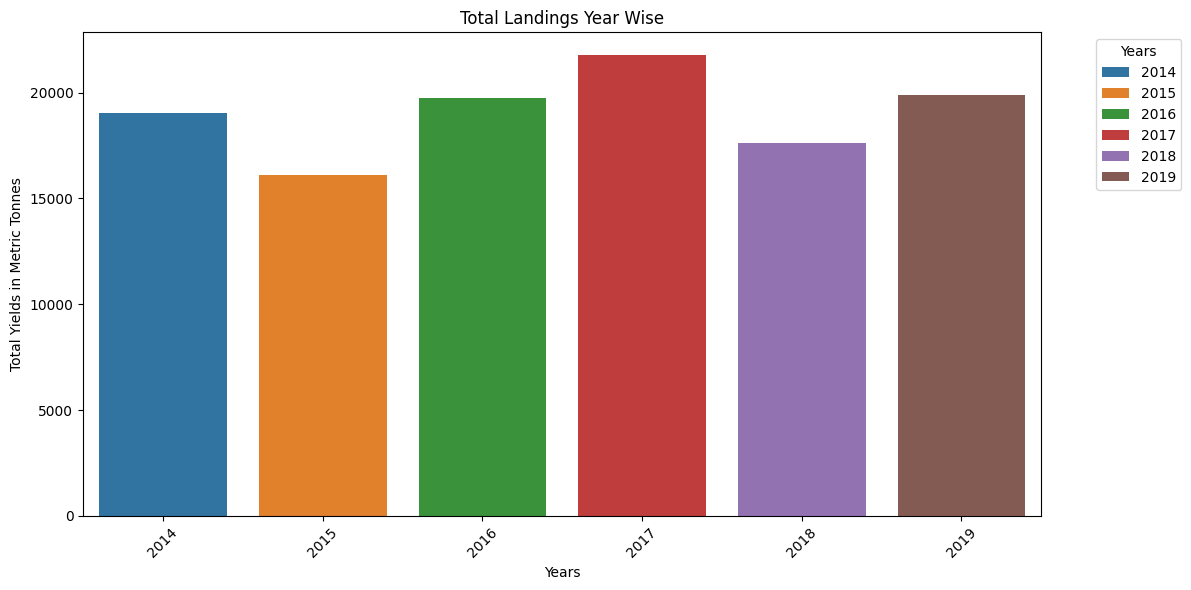
if \_\_name\_\_ == "\_\_main\_\_":

main()

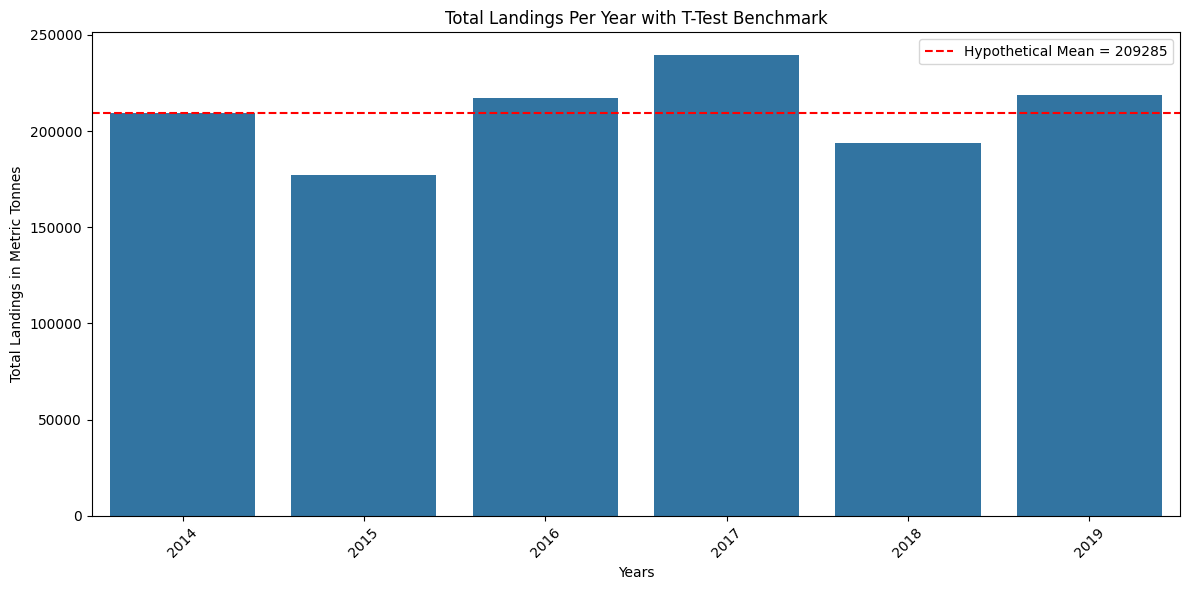
Correlation Analysis Summary:  
  
Trawlnet:  
 Strongest Positive Correlation: Non-mechanised (0.741)  
 Strongest Negative Correlation: Others (-0.631)  
  
Gillnet:  
 Strongest Positive Correlation: Non-mechanised (0.229)  
 Strongest Negative Correlation: Hook & Line (-0.844)  
  
Hook & Line:  
 Strongest Positive Correlation: Ringseine (0.697)  
 Strongest Negative Correlation: Gillnet (-0.844)  
  
Ringseine:  
 Strongest Positive Correlation: Hook & Line (0.697)  
 Strongest Negative Correlation: Gillnet (-0.789)  
  
Non-mechanised:  
 Strongest Positive Correlation: Trawlnet (0.741)  
 Strongest Negative Correlation: Others (-0.659)  
  
Others:  
 Strongest Positive Correlation: Gillnet (0.152)  
 Strongest Negative Correlation: Non-mechanised (-0.659)

**Graphs**  
**IJ ROY**

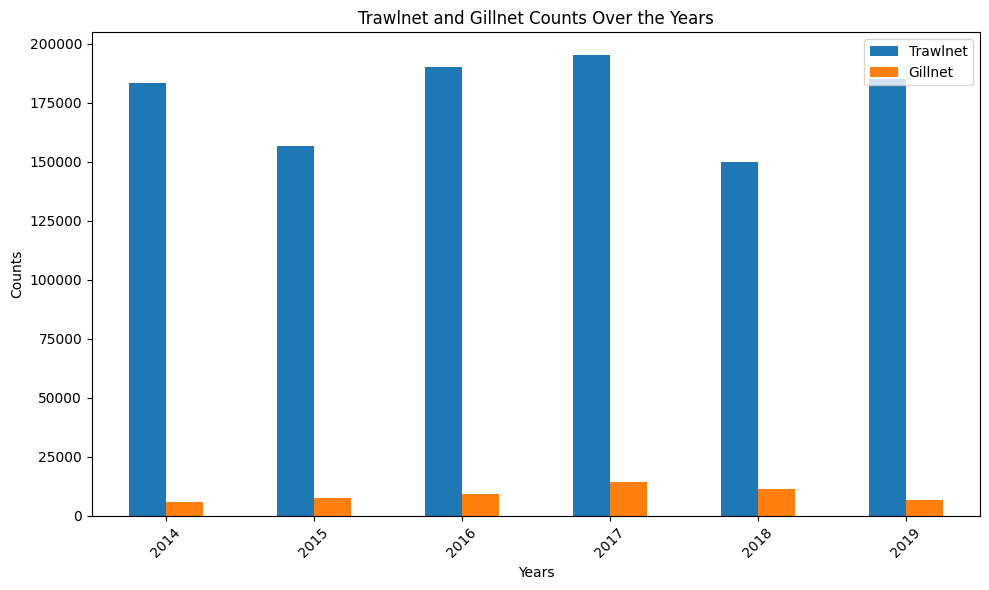
**Graph 1:**

This bar plot provides a visual summary of the annual landings of *T. lepturus* in the maritime states along the Indian EEZ during 2014–2019, highlighting trends, variations, and comparative yields over the years in metric tonnes.

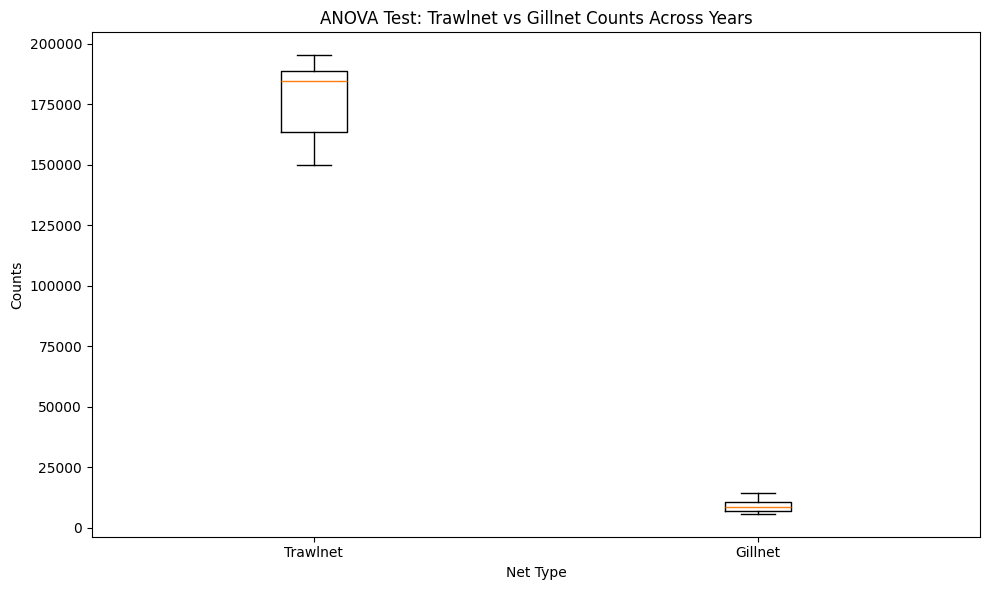
**Graph 2:**

The bar plot provides a visual summary of the total landings per year, along with a benchmark line indicating the hypothetical mean. This highlights variations in yearly landings and allows for comparison against the mean, aiding in the interpretation of potential trends and deviations.

**Graph 3:**

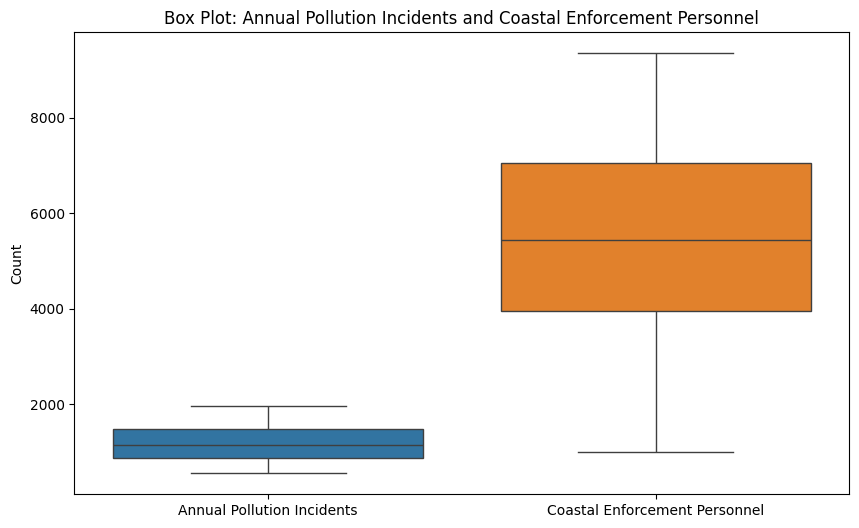
The bar plot provides a visual summary of the annual *T. lepturus* landings in trawl nets and gillnets for total of each maritime state of the Indian EEZ. It highlights the distribution of counts over the years, allowing for easy comparison between different net types.

**Graph 4:**

The box plot provides a visual summary for the ANOVA test, comparing *T. lepturus* counts between trawlnets and gillnets across years. It highlights the median, quartiles, and potential outliers for each net type, aiding in the assessment of variability and differences in their distributions.

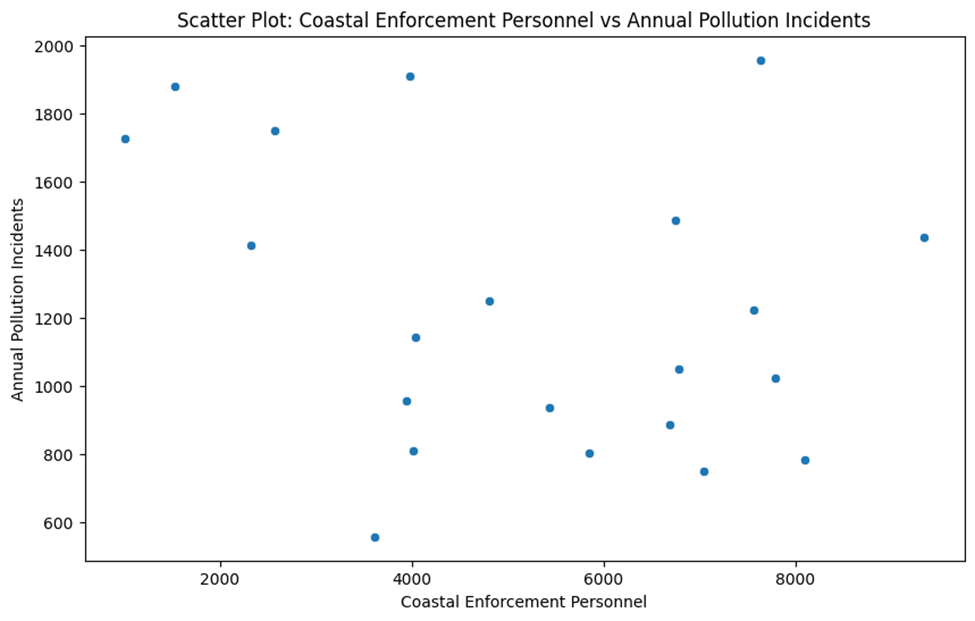
SARVESH GULGULIA

GRAPH 1



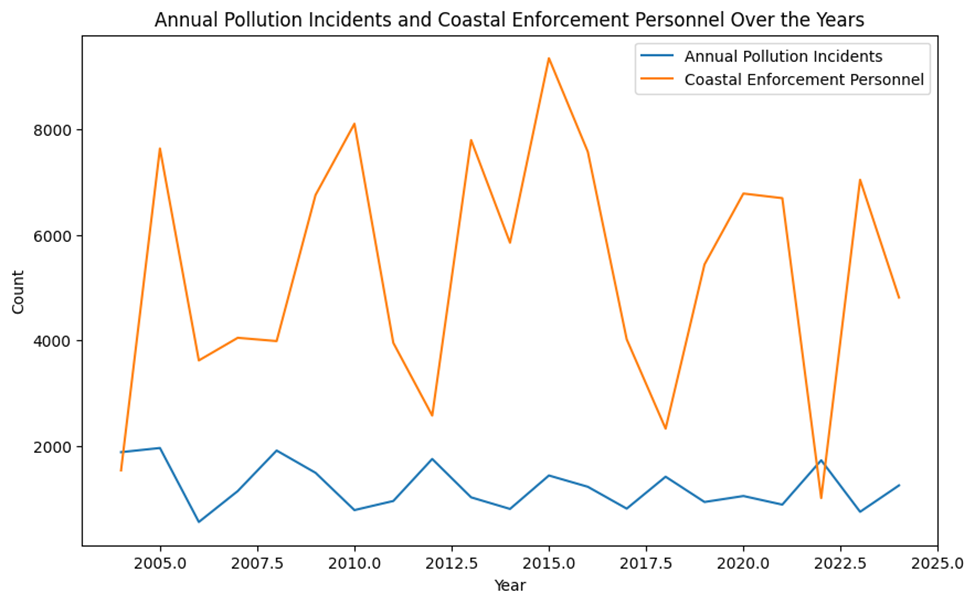
The box plot provides a visual summary of the distribution of Annual Pollution Incidents and Coastal Enforcement Personnel. It highlights the median, quartiles, and potential outliers in the data.

GRAPH 2



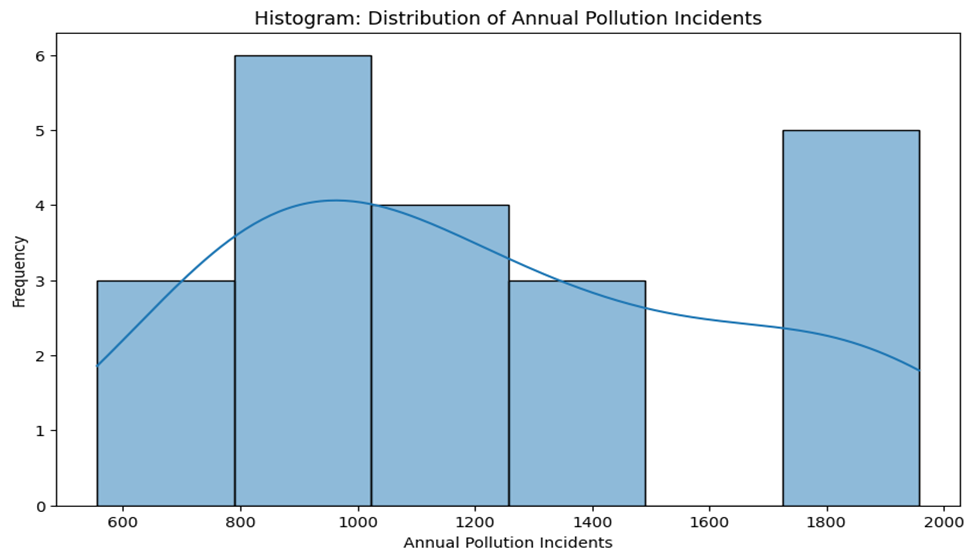
The scatter plot visualizes the relationship between Coastal Enforcement Personnel and Annual Pollution Incidents. It helps identify any patterns or correlations between the two variables.

GRAPH 3



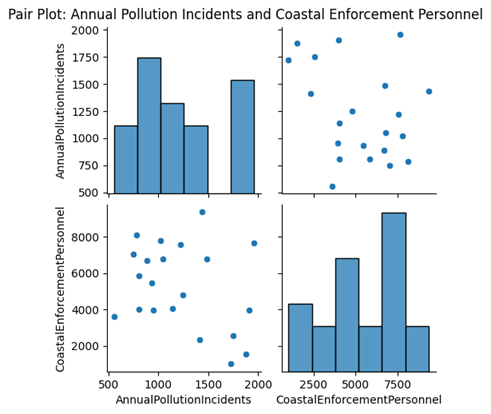
The line plot shows fluctuations in both Annual Pollution Incidents and Coastal Enforcement Personnel over the years. There is no clear trend indicating that increasing personnel consistently reduces pollution incidents, suggesting that other factors may also play a significant role.

GRAPH 4



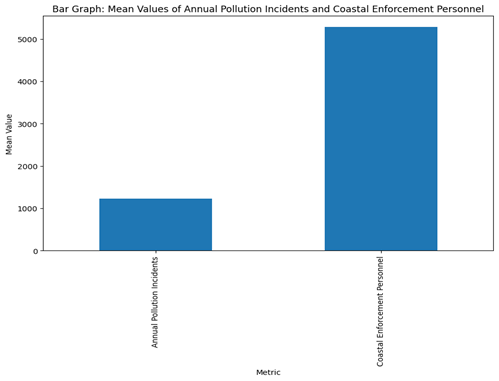
The histogram provides a visual representation of the distribution of Annual Pollution Incidents. It shows the frequency of different incident counts.

Graph 5



The pair plot shows scatter plots and histograms for each pair of variables. It indicates a weak negative correlation between Coastal Enforcement Personnel and Annual Pollution Incidents, consistent with the scatter plot analysis.

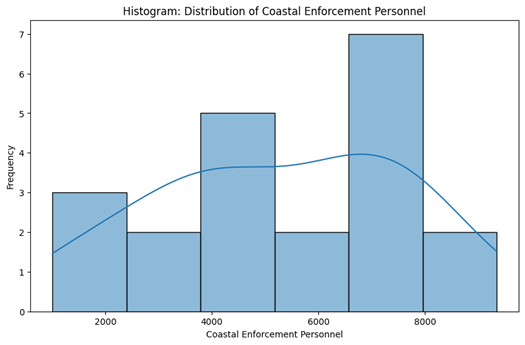
Graph 6



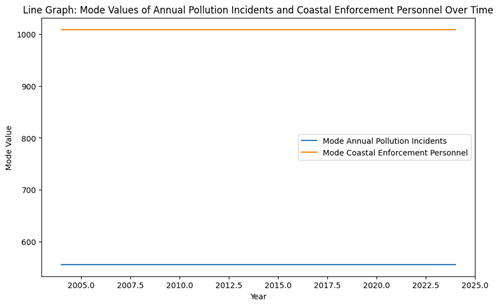
The bar graph shows that the mean value of Coastal Enforcement Personnel is significantly higher than that of Annual Pollution Incidents. This highlights the difference in scale between the two metrics.

Graph 7

The histogram shows that the distribution of Coastal Enforcement Personnel is relatively symmetrical with a slight negative skew. The presence of a few high-frequency bins indicates that certain personnel counts are more common.

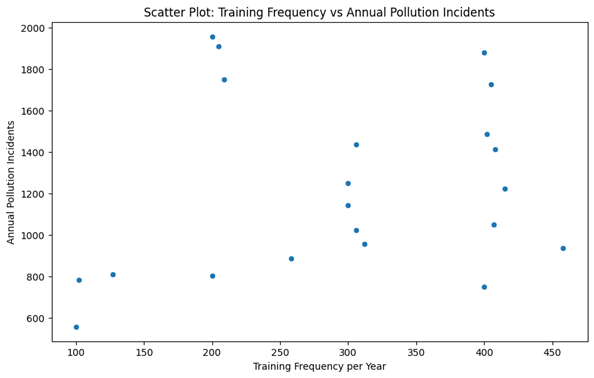


GRAPH 8



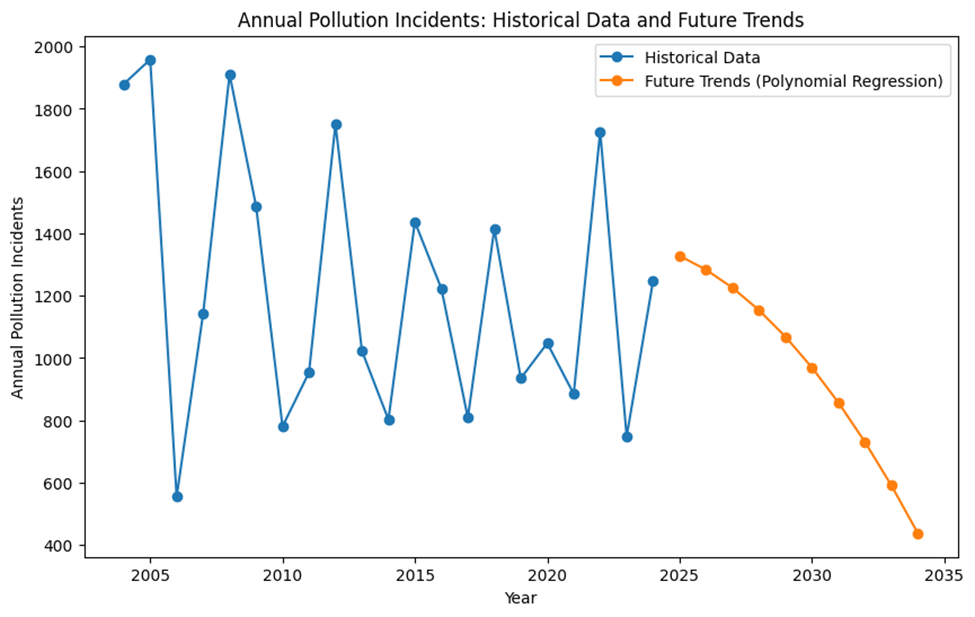
The line graph shows that the mode values for both Annual Pollution Incidents and Coastal Enforcement Personnel remain relatively constant over time. This suggests that the most common values do not change significantly from year to year.

GRAPH 9



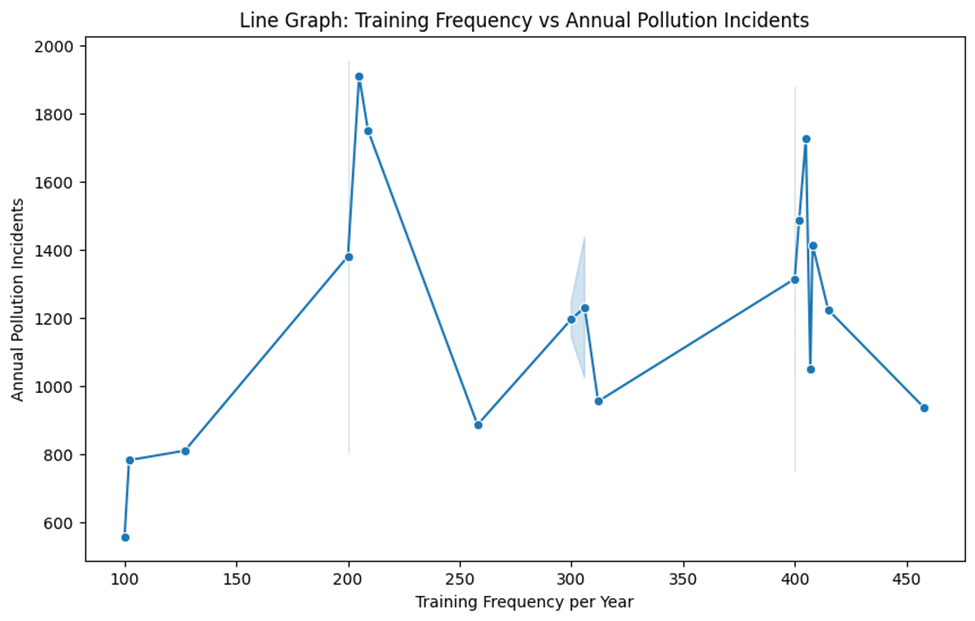
The scatter plot shows a weak positive correlation between Training Frequency and Annual Pollution Incidents. The data points are scattered without a clear trend, supporting the correlation result. This visual representation confirms that training frequency does not significantly influence the number of pollution incidents.

GRAPH 10

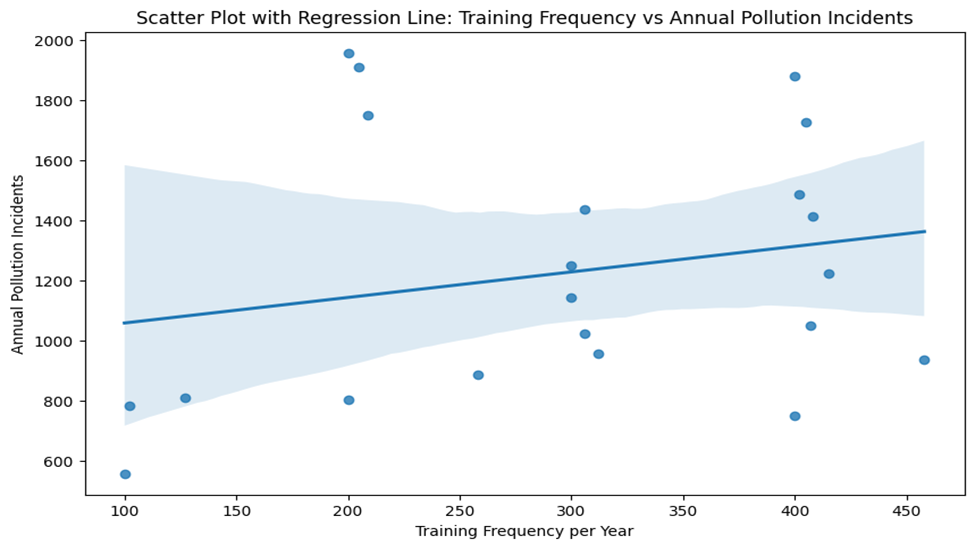


The line plot shows a decreasing trend in Annual Pollution Incidents over the next decade, as predicted by the polynomial regression model. This suggests that the factors included in the model (Marine Pollution Budget Ratio, Personnel per km of Coast, Pollution Response Vessels) will contribute to a reduction in pollution incidents over time.

GRAPH 11



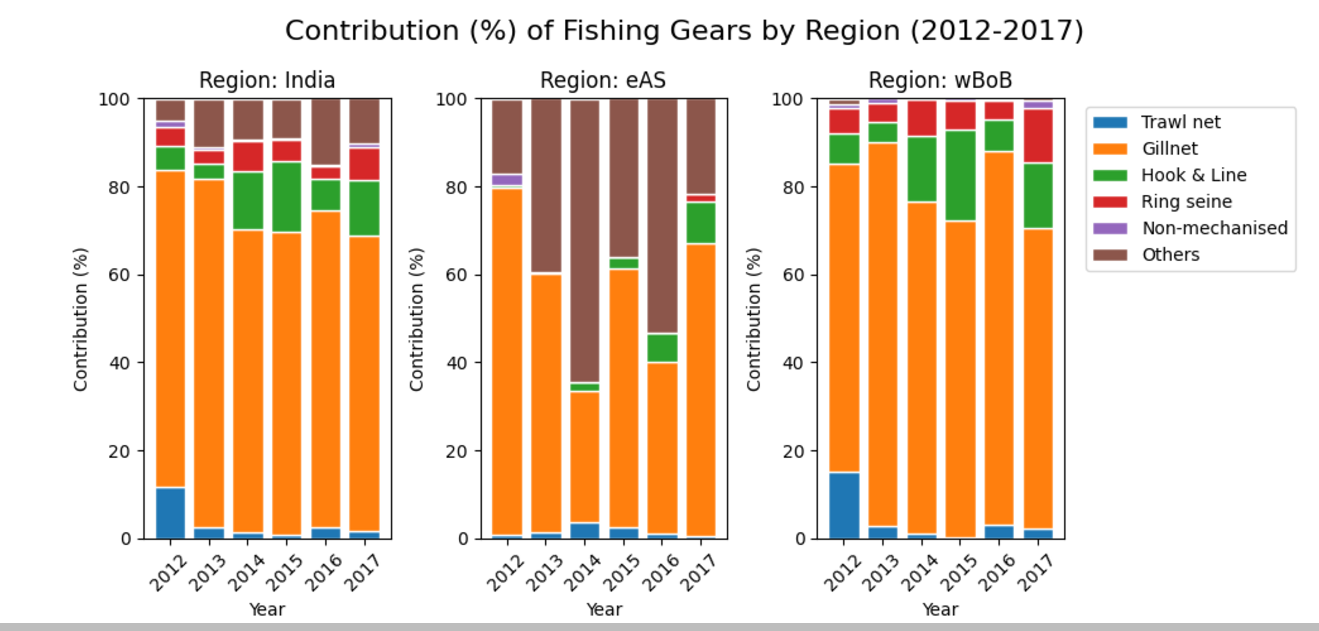
The line graph shows no clear trend indicating that changes in training frequency significantly impact Annual Pollution Incidents. The data points are scattered without a consistent pattern, supporting the ANOVA result.



The scatter plot with a regression line shows a weak negative correlation between Training Frequency and Annual Pollution Incidents. The regression line is relatively flat, indicating that changes in training frequency have little to no impact on pollution incidents. This visual representation aligns with the ANOVA result, confirming that training frequency does not significantly influence the number of pollution incidents.

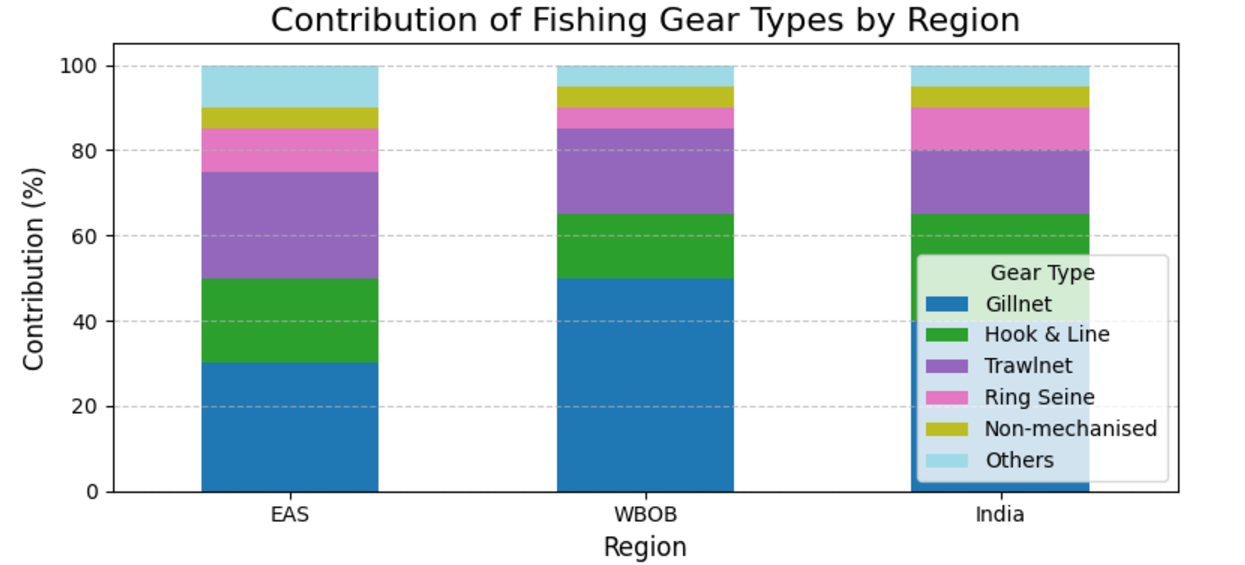
**SNEHA MISHRA**

GRAPH 1:

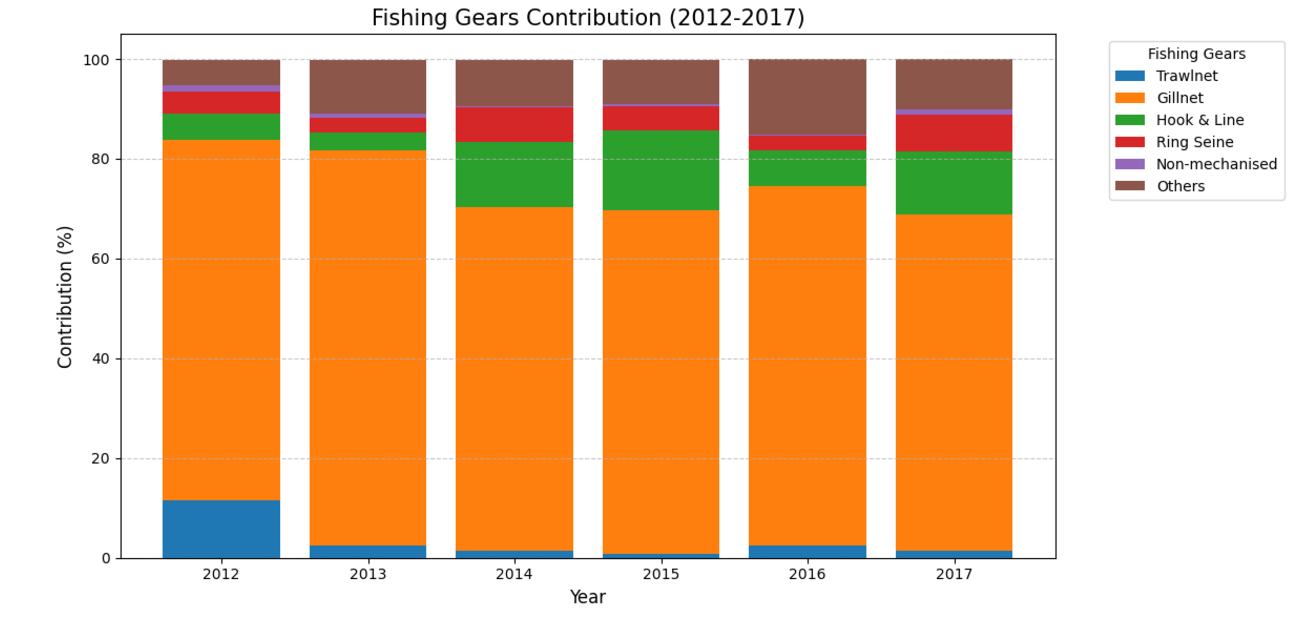


The graph analyzes the contribution (percentage) of different fishing gears used in three regions - India, eAS, and WBoB - from 2012 to 2017. It shows the relative usage and trends of various fishing technologies like trawl net, gillnet, hook & line, ring seine, and non-mechanised gears in these regions over the given time period.

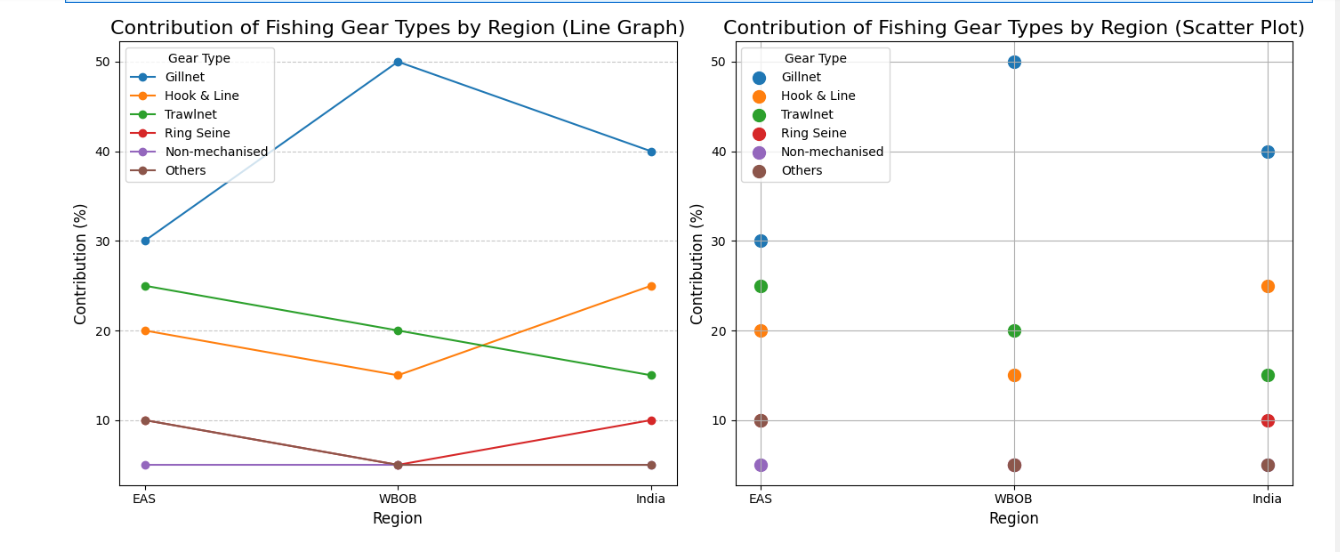
GRAPH 2:



GRAPH 3:



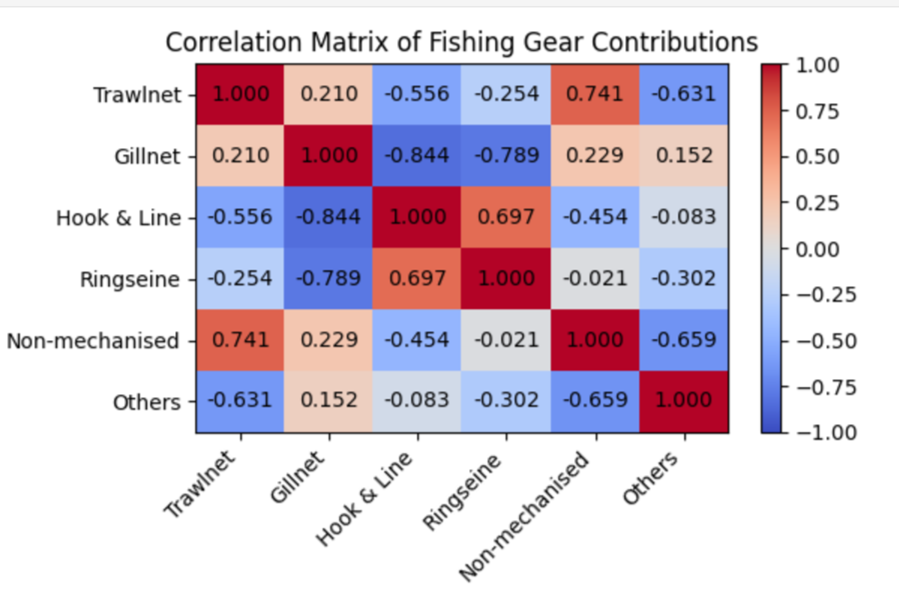
GRAPH 4:



The graphs present a detailed analysis of the contribution (percentage) of different fishing gear types across three regions - EAS, WBoB, and India - over the years 2012 to 2017. The line graph shows the trend lines for each gear type, while the scatter plot provides a more granular view of the relative usage of the gear types in each region.

This data can help policymakers and fisheries managers understand the current fishing practices, identify opportunities for promoting more sustainable gear usage, and track progress towards their management goals over time.

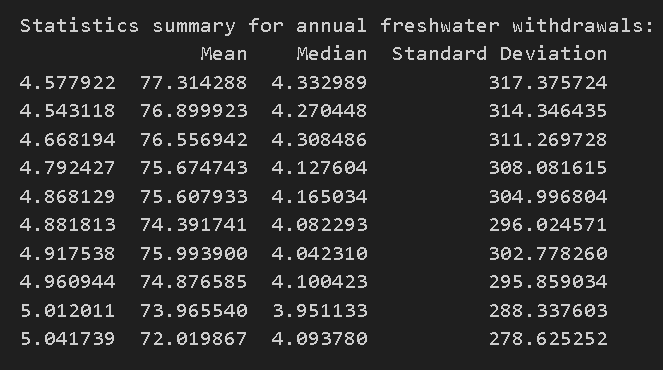
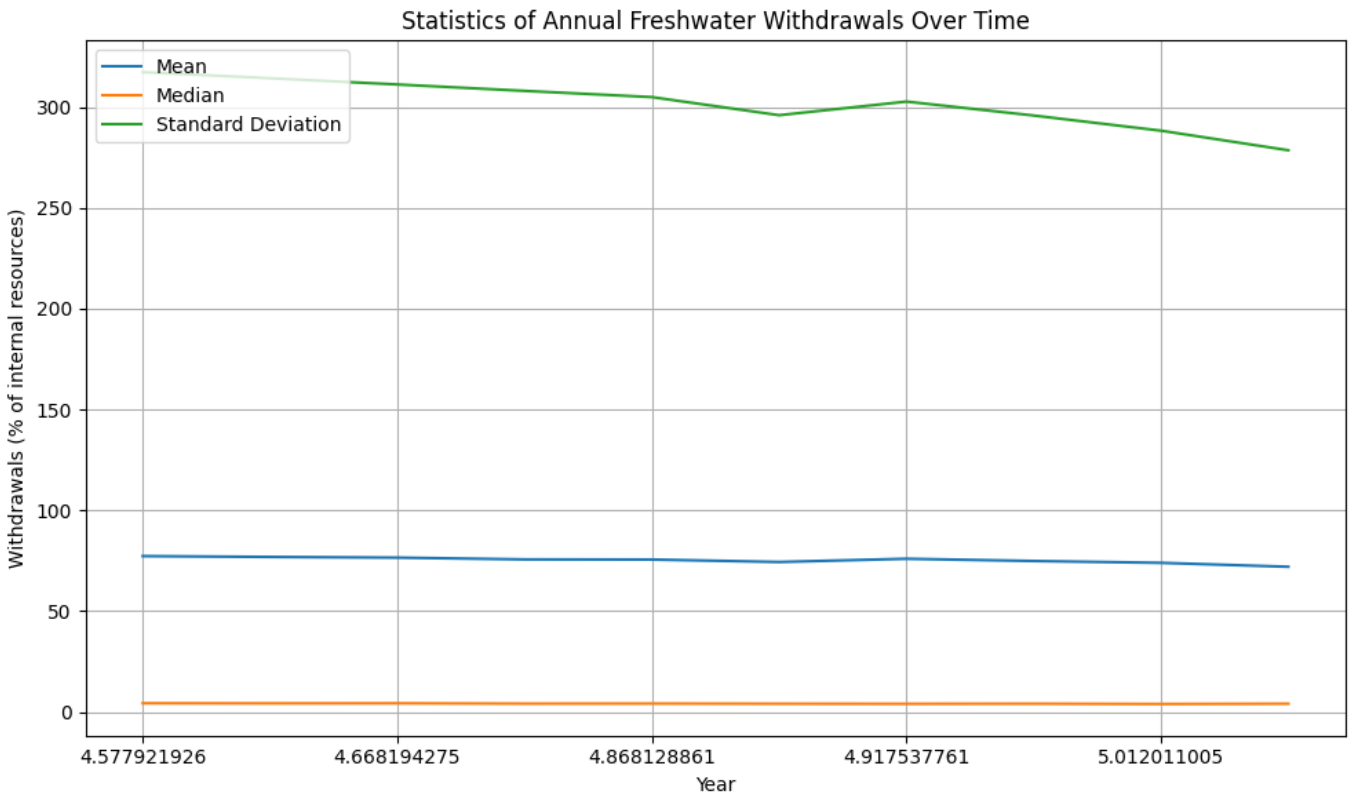
GRAPH 5:

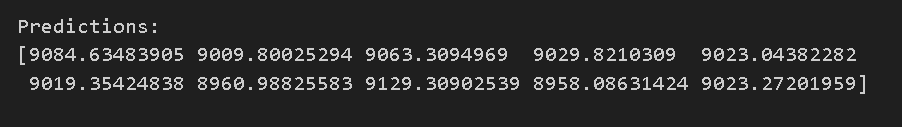
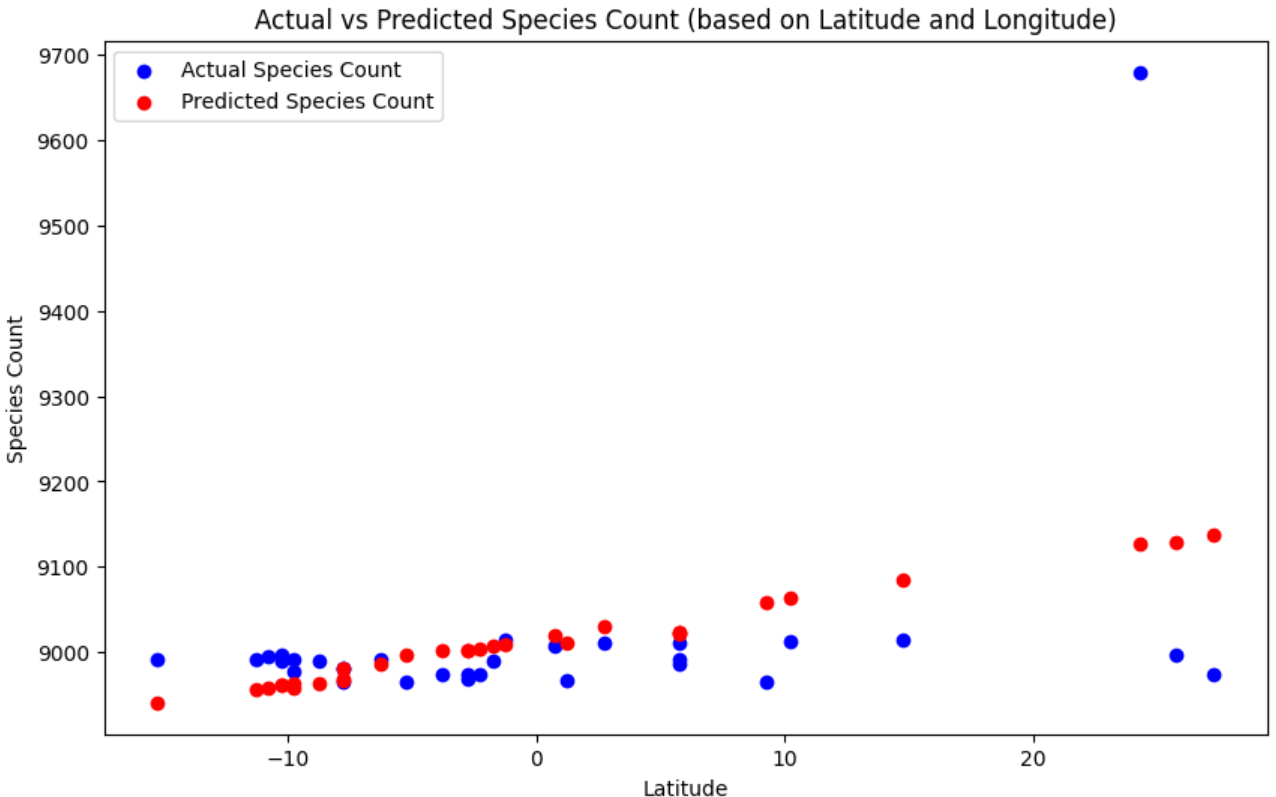


The graph presents the correlation matrix of fishing gear contributions across different gear types. It shows the strength and direction of the relationships between the usage of various fishing gears, such as trawlnet, gillnet, hook & line, ringseine, non-mechanised, and others. The correlation values range from -1 to 1, indicating the degree of positive or negative correlation between the gear types.

SHUBHAM SHARMA

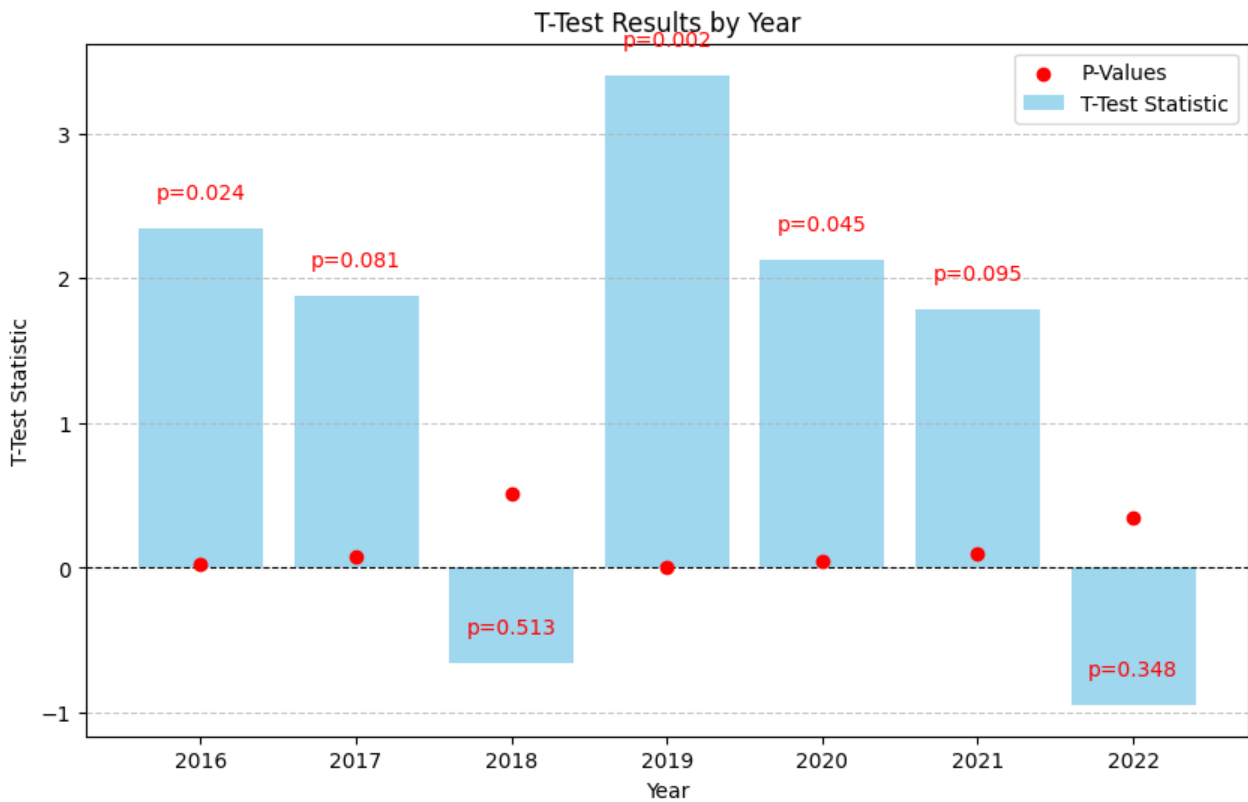
GRAPH 1:

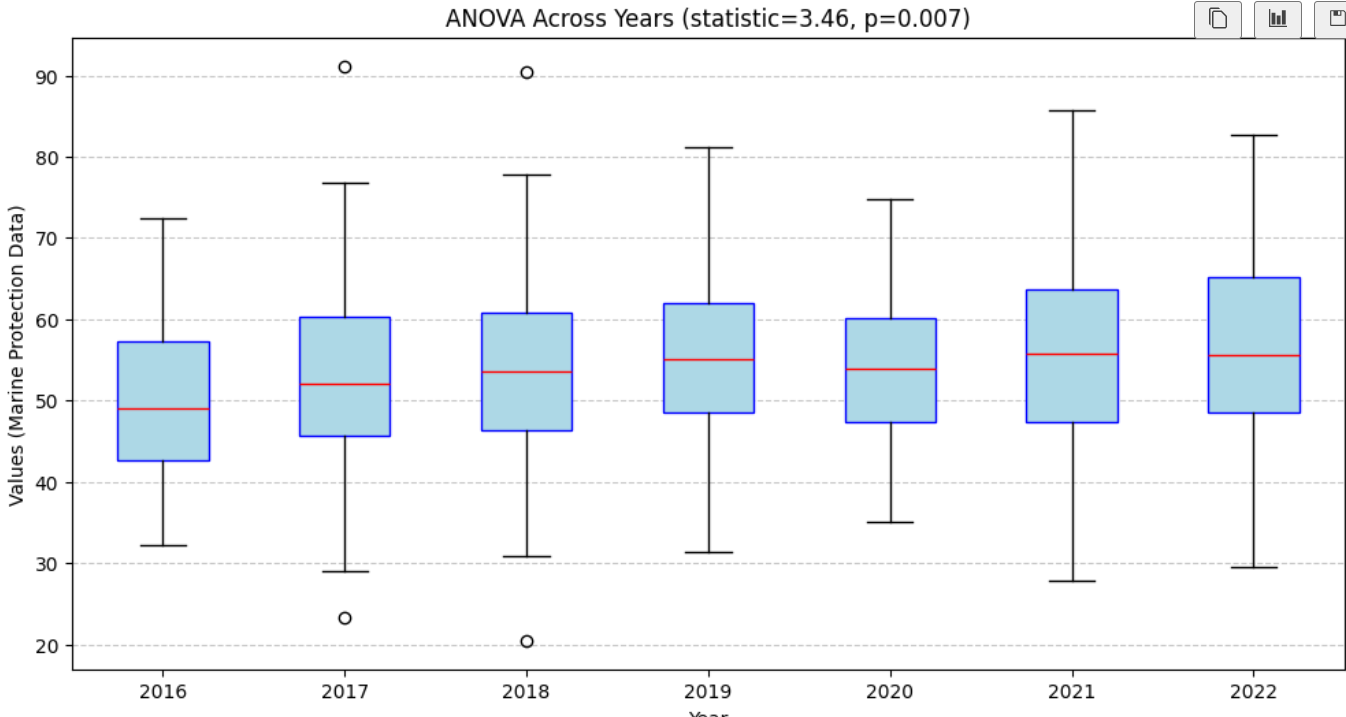






t-test





**Conclusion**

The report emphasizes the issues and potential to achieve the Sustainable Development Goal 14 Life Below Water as it pertains to the Indian Exclusive Economic Zone. The statistical and empirical investigations led also to the derive of the following key insights:

Marine Pollution Control: It has been established by the linear and polynomial regression analysis that persons per kilometer of coastline are the major factors affecting incidents of pollution, with minimum effect posed by budgetary provisions and vessels for burning cases of pollution.

Marine Biodiversity: The biodiversity varied due to different degree and zones of freshwater runoffs and human impacts, hence, the need for concentrated conservation.

Sustainable Fishing Practices Seem Necessary: Gear type has a major impact on the yield as well as sustainability of fisheries. Findings from ANOVA and T-test indicated that in as much as the gear may have been determined; it should be optimized in respect of ecological, economic and social set objectives.

Effectiveness of Training Programs and Policy: In as much as the occurrence of pollution incidences was understated due to frequency of training programs, it was established that regular monitoring, enforcement and stakeholder participation were significant variables.