

# Detailed Computational Project-1

## Applied Statistics - MA4240

Sai Kaushik .P - AI23BTECH11017

Pranav JVS - AI23BTECH11009

Samagalla Suresh Kumar - AI23BTECH11024

## 1 Introduction

The Data file used for this project was `Quality_of_Life.csv` . It's Source being ([Source](#)) from [www.kaggle.com](http://www.kaggle.com) . The data is measure of different aspects of possible/ important life supporting features the countries(236 in total) hold.

## 2 Observations from the CSV File

### 2.1 Dataset Overview

The dataset contains 236 countries with 11 columns. It includes both numerical values (e.g., Purchasing Power Value, Safety Value) and categorical labels (e.g., Purchasing Power Category, Safety Category).

#### Key Attributes:

- Purchasing Power: Measured numerically and classified into categories (Very Low, Low, Moderate, High, Very High).
- Safety: A numerical score with a corresponding categorical classification.
- Health Care: Includes a numerical score and a category (Very Low, Low, Moderate, High, Very High).
- Cost of Living: Provides a numerical value and a categorical classification.
- Pollution: A numeric pollution score along with its category.

#### Notable Data Points:

- Aland Islands has one of the highest purchasing power values (125.01) and a high safety score (71.81).
- Alderney has extreme values, with a perfect health care score (100.00) and the lowest pollution (1.72).
- Afghanistan has low scores in purchasing power (32.15) and safety (25.33), indicating economic and security challenges.
- Algeria has a moderate safety value (47.54) but a high pollution score (63.87).

#### Possible Issues:

- Some countries, like "Alderney," have extreme values (e.g., 0.00 for cost of living).

### 3 Observations from the (.ipynb) Execution

#### Purchasing Power Analysis:

- Europe has the highest average purchasing power (91.61), with the highest third quartile and median values.
- Asia follows with an average of 68.49, while North America is lower at 51.02.
- Africa and South America have significantly lower purchasing power, around 32–33.
- Oceania's data is inconsistent, with a median of 0.00, likely due to missing or extreme values.
- Antarctica has zero values, indicating a lack of data or economic activity.

#### Safety Index:

- Antarctica reports the highest average safety (71.18), which may not be meaningful due to limited data.
- Europe is the safest region overall (66.03), followed by Asia (60.97).
- Africa (45.73) and South America (38.31) have lower safety indices, confirming known security challenges.
- North America's safety average is 51.42, which is lower than Asia and Europe.

#### Health Care Scores:

- Europe leads in healthcare (66.54 average), followed by Asia (59.47).
- North and South America have similar healthcare averages (52-53).
- Oceania and Antarctica have large variations, likely due to missing or inconsistent data.
- Africa has the lowest healthcare quality, with an average of 48.36.

#### Cost of Living Trends:

- Similar patterns as purchasing power—Europe has the highest cost of living, followed by Asia and North America.
- Africa and South America remain lower in cost, but their lower purchasing power means affordability remains a challenge.

#### Pollution Levels:

- Asia shows high pollution, which aligns with the main industrial activity.
- Europe has moderate levels of pollution, while Africa and South America have lower levels.
- Antarctica has minimal pollution, which is expected given its environment.

#### Conclusion:

- Europe consistently ranks high in purchasing power, healthcare, and safety.
- Africa and South America face economic and safety challenges, with low purchasing power and safety indices.
- Asia has strong purchasing power, but it struggles with pollution.
- North America has a mix of high and moderate values, with lower safety compared to Asia and Europe.
- Antarctica's data is unreliable, possibly due to missing values or the nature of its population.

## 4 OutPuts from code

### 4.1 For Purchasing Powers

#### For Africa's Purchasing Power Values

The average is 32.530  
The median is 21.435  
Standard Dev is 40.645  
Mean dev is 25.461  
1st Quartile is 10.098  
2nd Quartile is 21.435  
3rd Quartile is 40.370  
range is 224.460

#### For Asia's Purchasing Power Values

The average is 68.489  
The median is 47.845  
Standard Dev is 53.036  
Mean dev is 42.329  
1st Quartile is 32.240  
2nd Quartile is 47.845  
3rd Quartile is 106.015  
range is 259.090

#### For Europe's Purchasing Power Values

The average is 91.613  
The median is 87.200  
Standard Dev is 41.950  
Mean dev is 33.344  
1st Quartile is 61.115  
2nd Quartile is 87.200  
3rd Quartile is 123.655  
range is 195.550

#### For North America's Purchasing Power Values

The average is 51.021  
The median is 44.295  
Standard Dev is 45.332  
Mean dev is 35.995  
1st Quartile is 8.197  
2nd Quartile is 44.295  
3rd Quartile is 76.833  
range is 153.020

#### For South America's Purchasing Power Values

The average is 33.377  
The median is 39.760  
Standard Dev is 17.888  
Mean dev is 14.096  
1st Quartile is 23.547  
2nd Quartile is 39.760  
3rd Quartile is 42.812

range is 57.920

#### For Oceania's Purchasing Power Values

The average is 43.418  
The median is 0.000  
Standard Dev is 82.918  
Mean dev is 60.813  
1st Quartile is 0.000  
2nd Quartile is 0.000  
3rd Quartile is 36.025  
range is 281.830

#### For Antarctica's Purchasing Power Values

The average is 0.000  
The median is 0.000  
Standard Dev is 0.000  
Mean dev is 0.000  
1st Quartile is 0.000  
2nd Quartile is 0.000  
3rd Quartile is 0.000  
range is 0.000

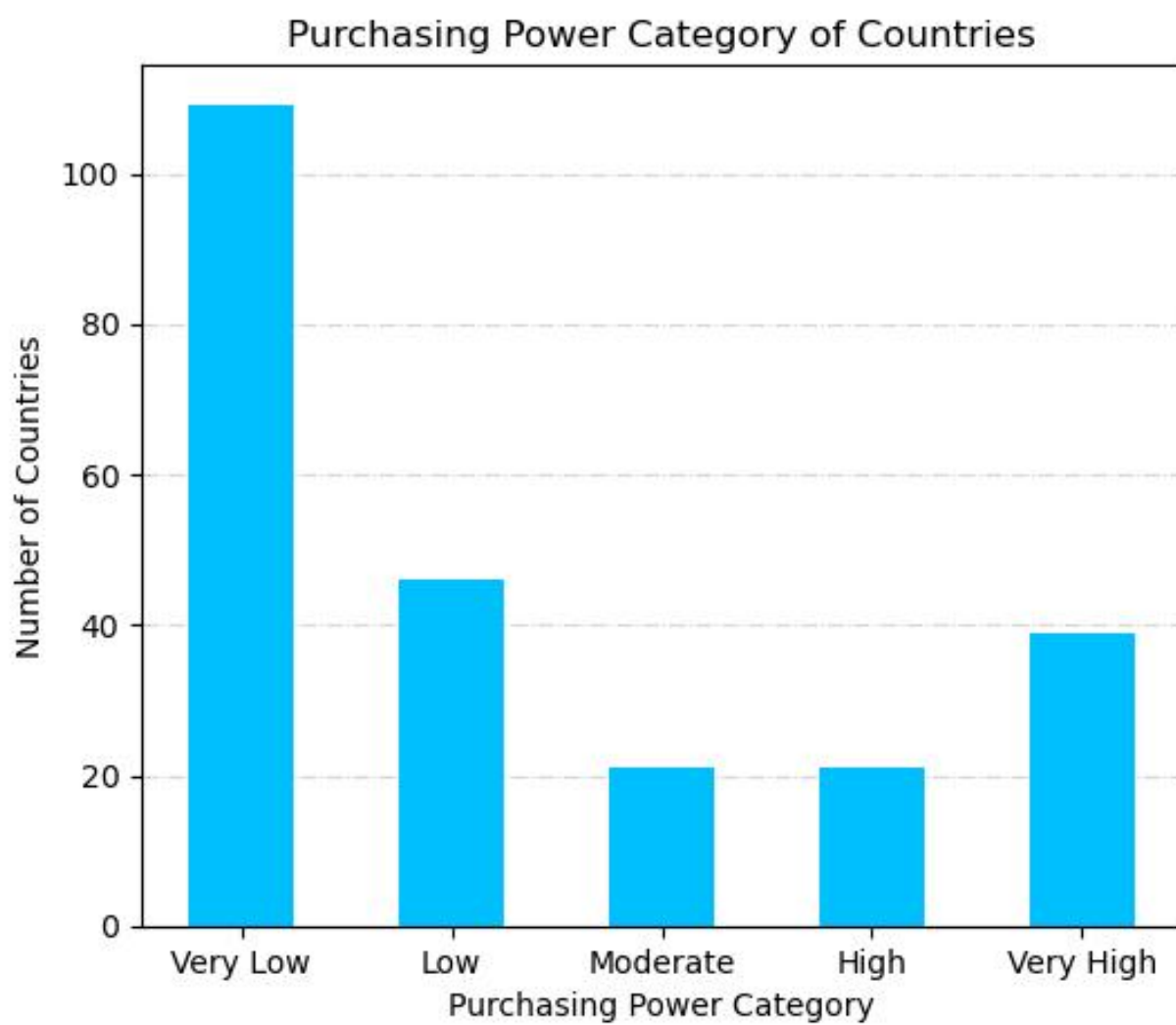


Figure 1: Bar Graph for Purchasing Power Categories.

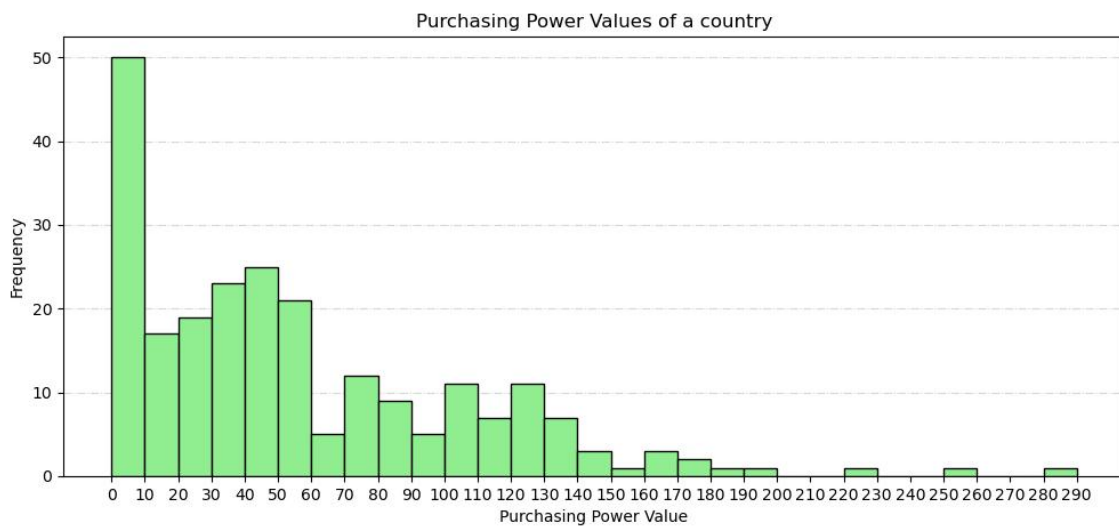


Figure 2: Histogram for Purchasing Power Values.

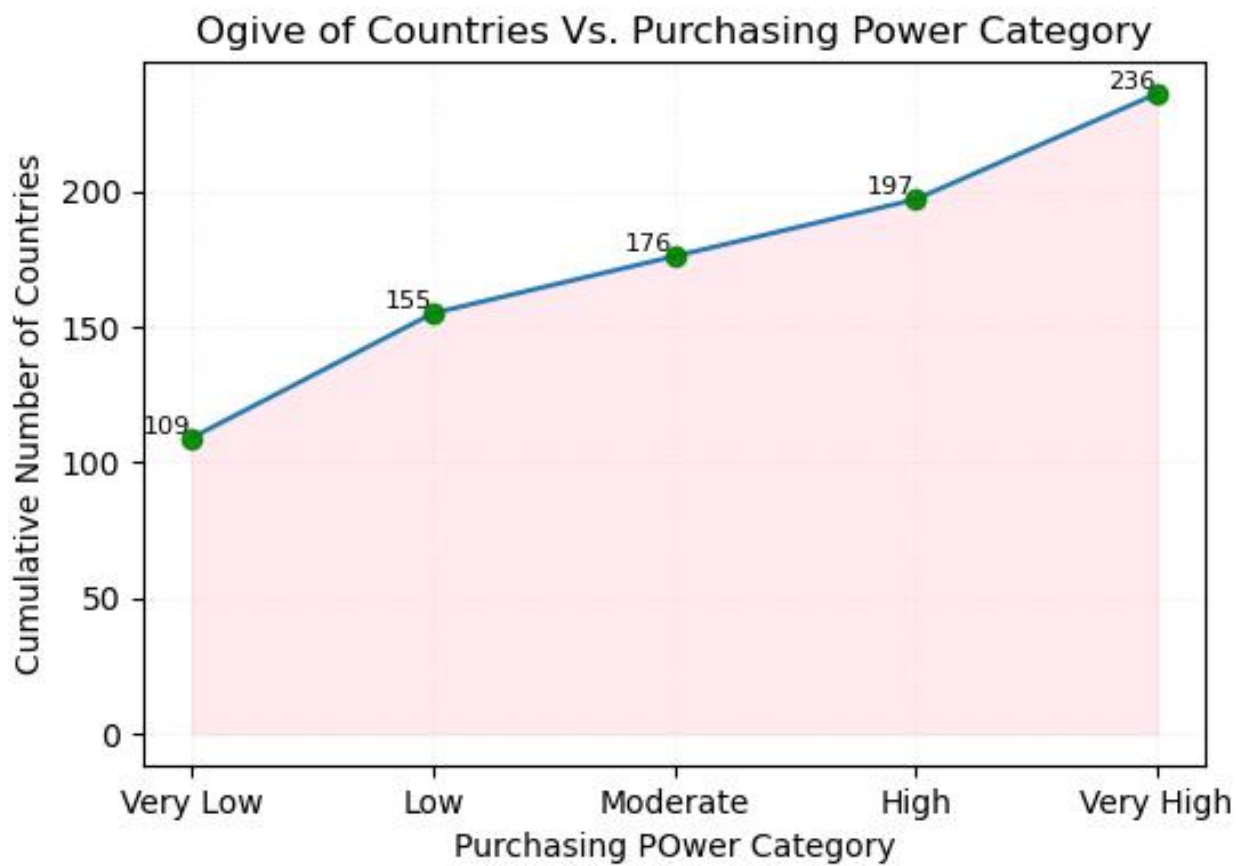


Figure 3: Ogive for Purchasing Power Categories.

Distribution of Countries by Purchasing Power Category

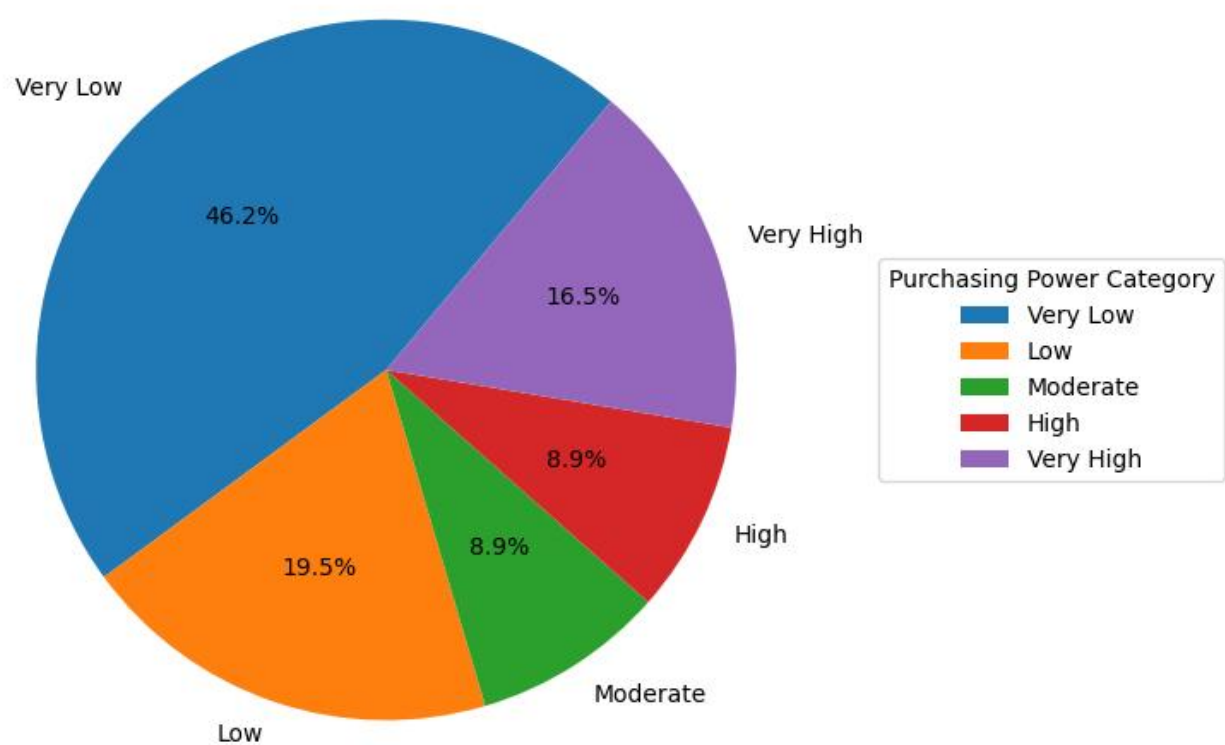


Figure 4: Pie Chart for Purchasing Power Categories.

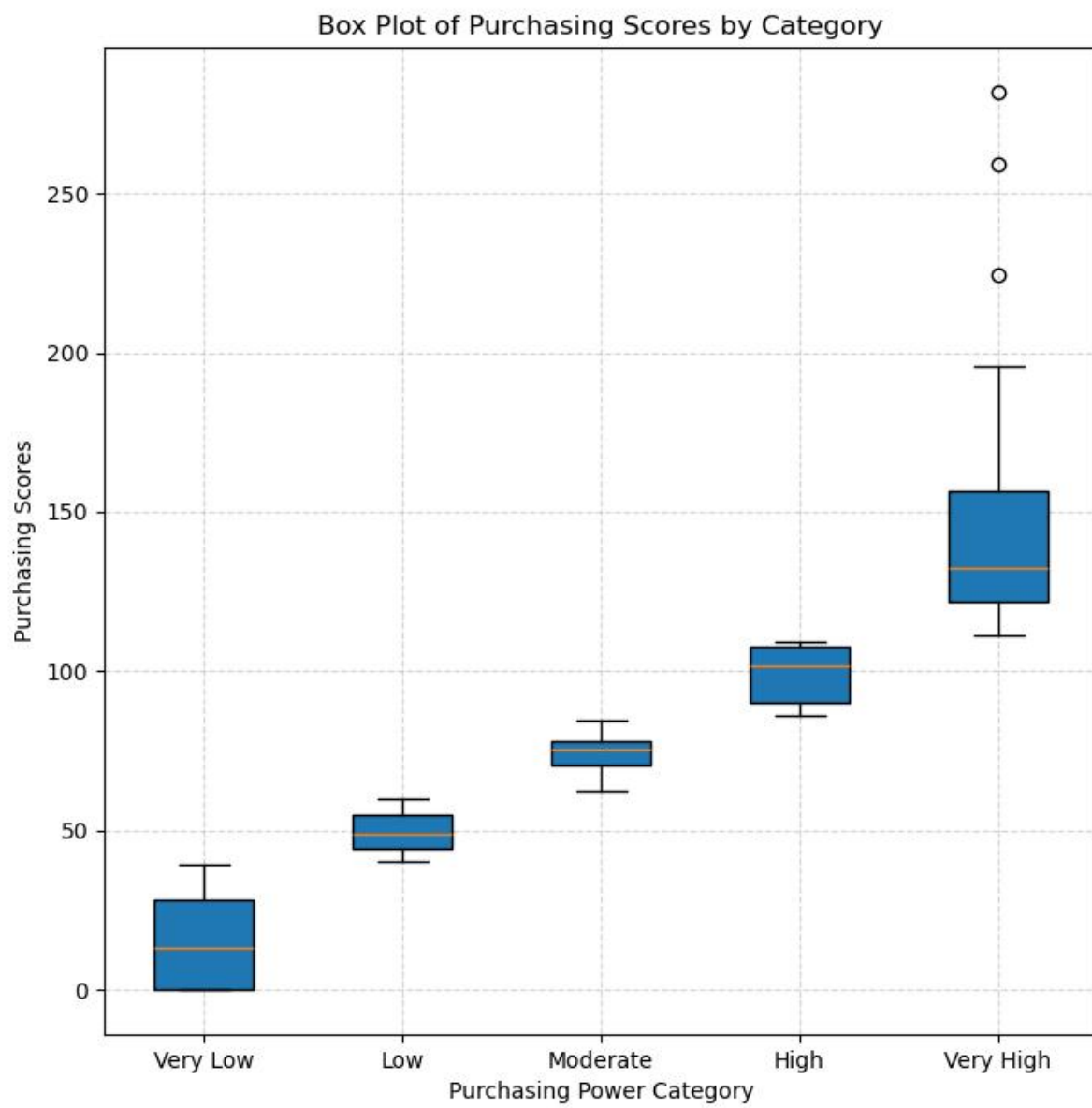


Figure 5: Box Plot for Purchasing Power Categories.



## Purchasing Power Values

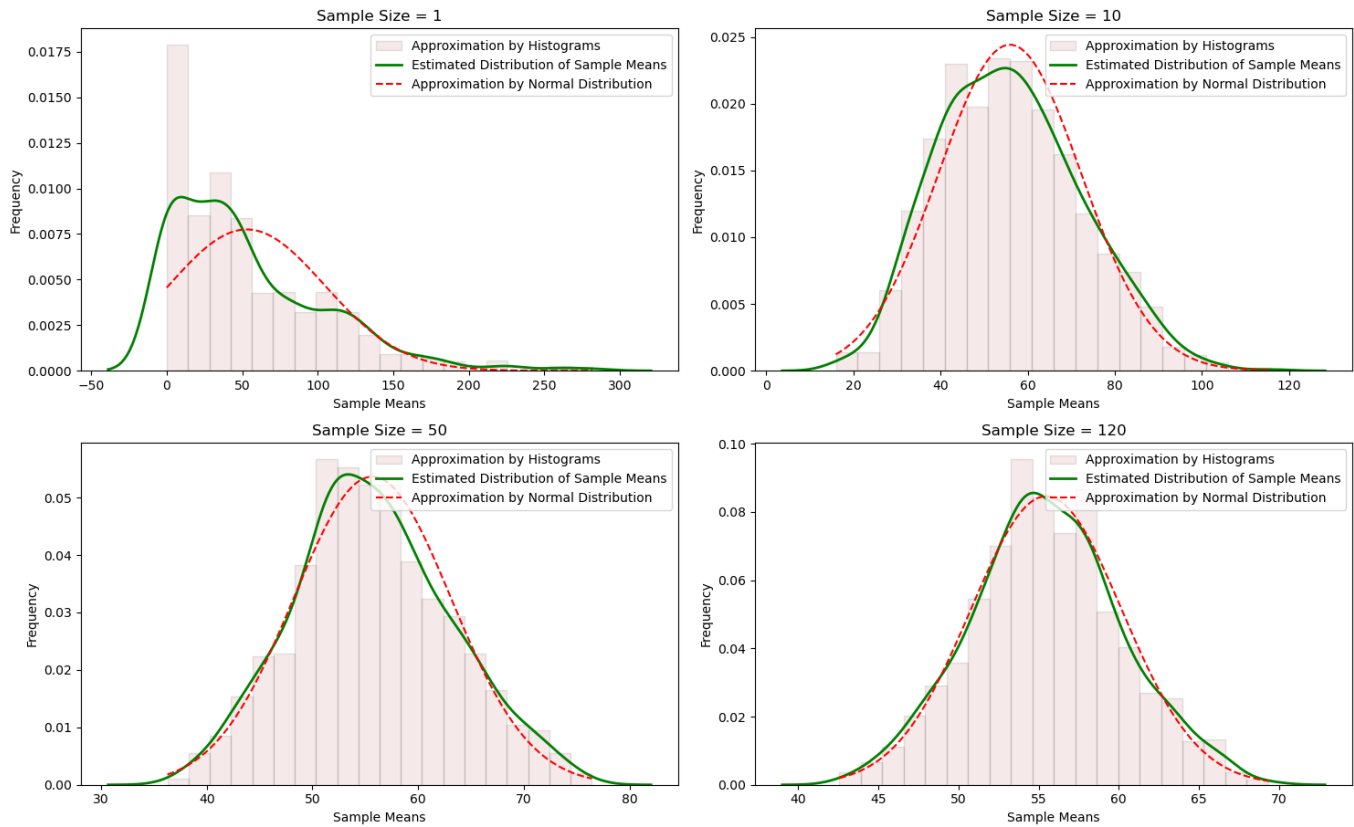


Figure 6: Approximation of Sample means for Purchasing Power values from CLT

## 4.2 For Safety

### For Africa's Safety Values

The average is 45.737  
 The median is 45.795  
 Standard Dev is 12.893  
 Mean dev is 9.744  
 1st Quartile is 36.850  
 2nd Quartile is 45.795  
 3rd Quartile is 52.935  
 range is 66.670

### For Asia's Safety Values

The average is 60.972  
 The median is 58.870  
 Standard Dev is 14.019  
 Mean dev is 11.468  
 1st Quartile is 51.720  
 2nd Quartile is 58.870  
 3rd Quartile is 72.353  
 range is 59.100

### For Europe's Safety Values

The average is 66.032  
 The median is 67.050

Standard Dev is 10.687  
Mean dev is 8.978  
1st Quartile is 57.695  
2nd Quartile is 67.050  
3rd Quartile is 74.035  
range is 43.280

#### For North America's Safety Values

The average is 51.428  
The median is 46.970  
Standard Dev is 17.175  
Mean dev is 13.070  
1st Quartile is 40.190  
2nd Quartile is 46.970  
3rd Quartile is 61.615  
range is 78.590

#### For South America's Safety Values

The average is 38.309  
The median is 37.220  
Standard Dev is 8.420  
Mean dev is 5.662  
1st Quartile is 35.090  
2nd Quartile is 37.220  
3rd Quartile is 39.767  
range is 36.380

#### For Oceania's Safety Values

The average is 50.921  
The median is 54.410  
Standard Dev is 22.687  
Mean dev is 17.175  
1st Quartile is 37.965  
2nd Quartile is 54.410  
3rd Quartile is 66.555  
range is 81.860

#### For Antarctica's Safety Values

The average is 71.188  
The median is 62.950  
Standard Dev is 16.302  
Mean dev is 13.626  
1st Quartile is 62.870  
2nd Quartile is 62.950  
3rd Quartile is 83.790  
range is 38.970

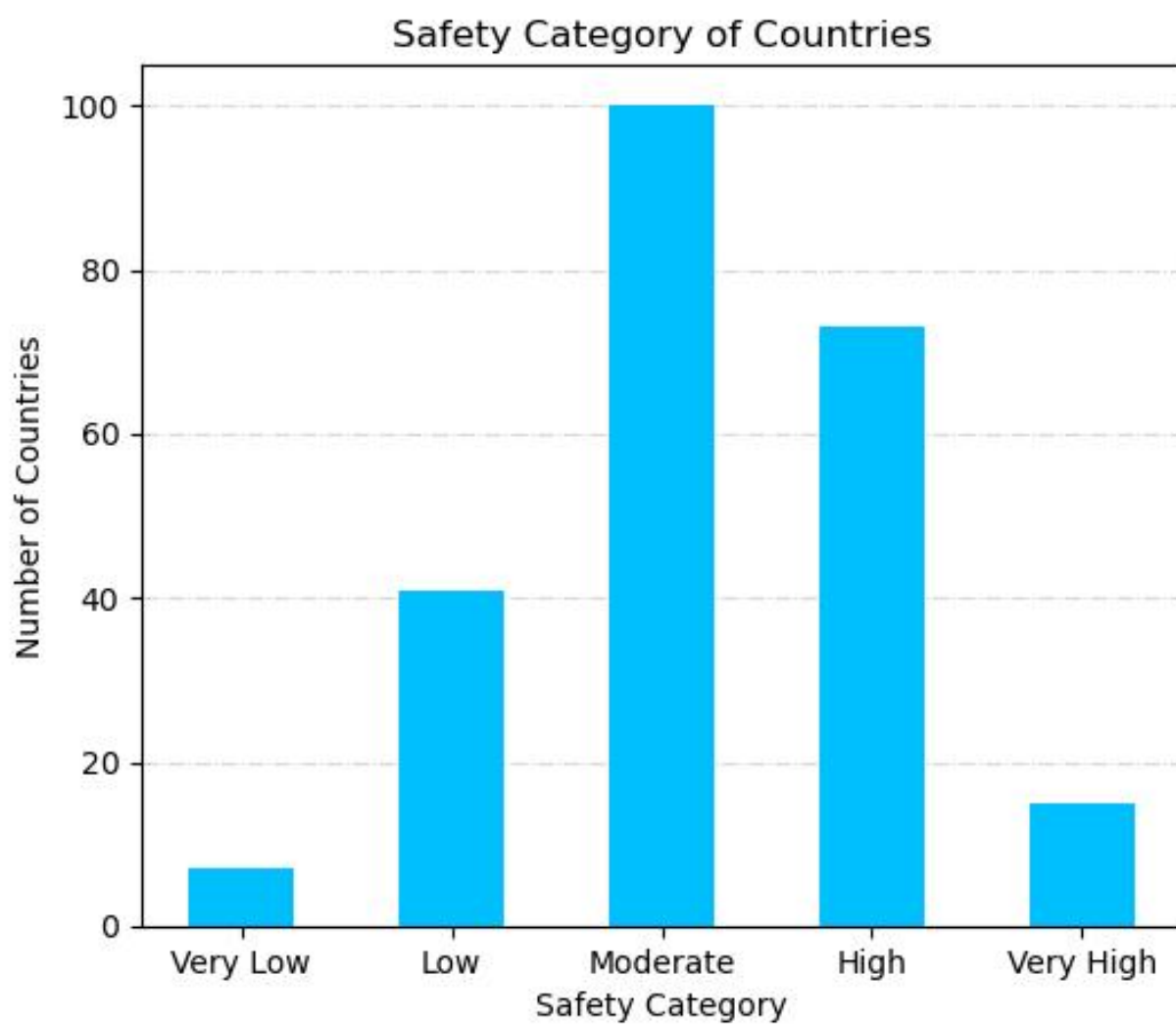


Figure 7: Bar Graph for Safety Categories.

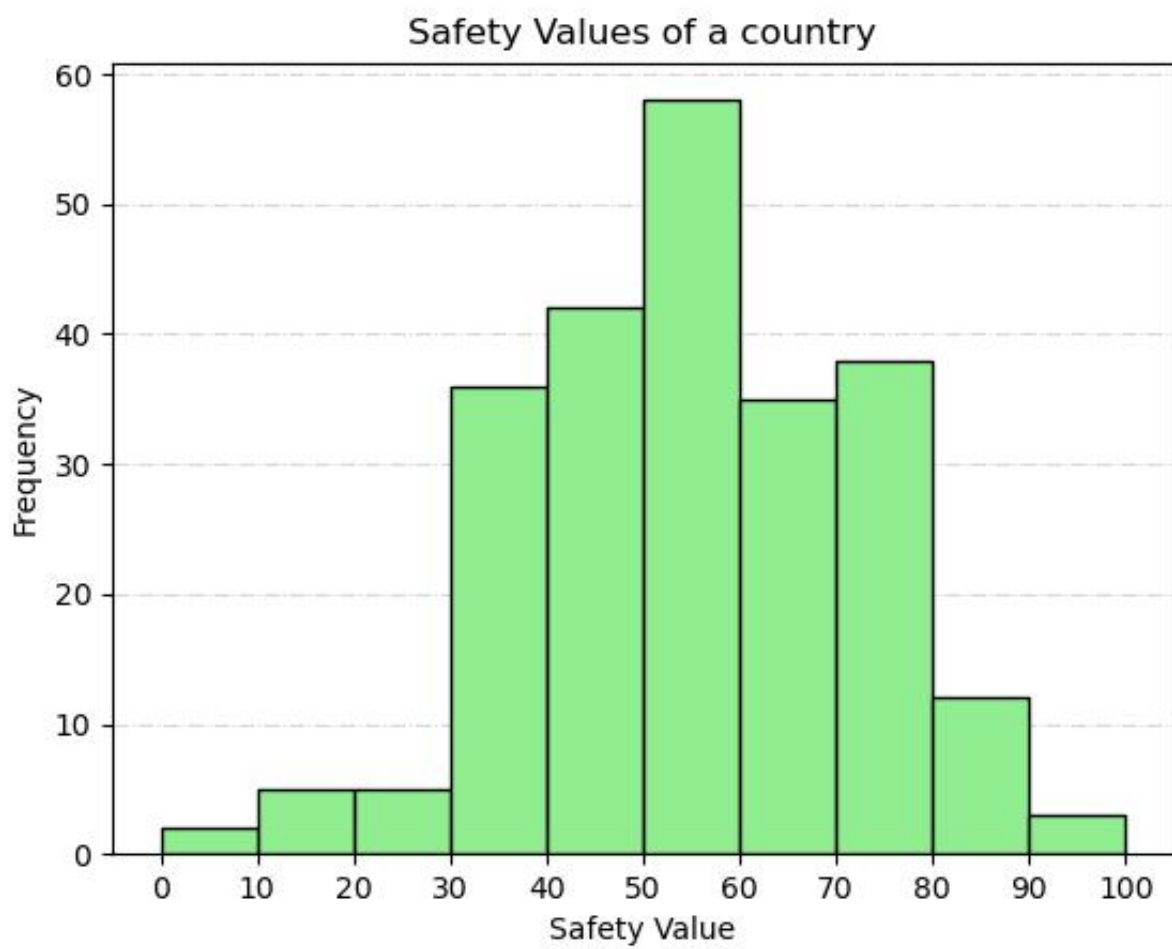


Figure 8: Histogram for Safety Values.

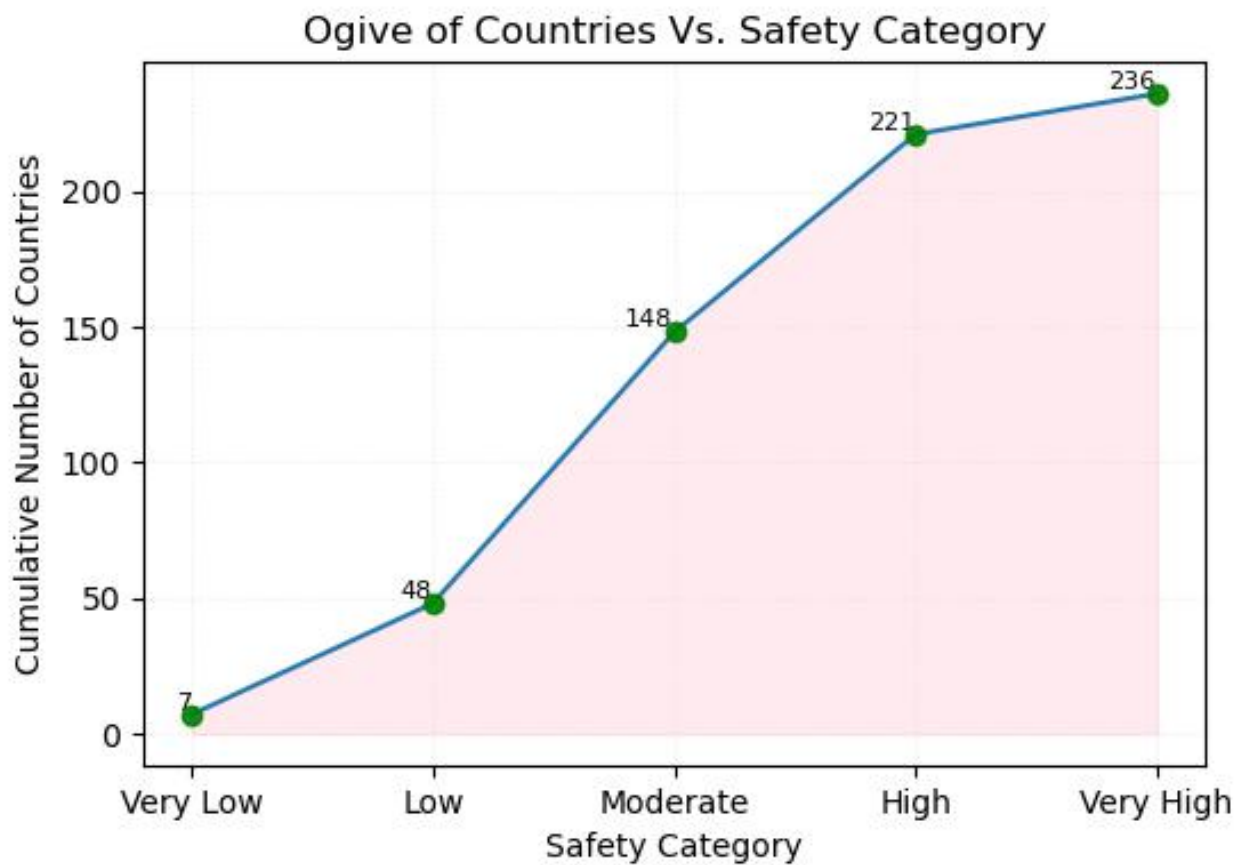


Figure 9: Ogive for Safety Categories.

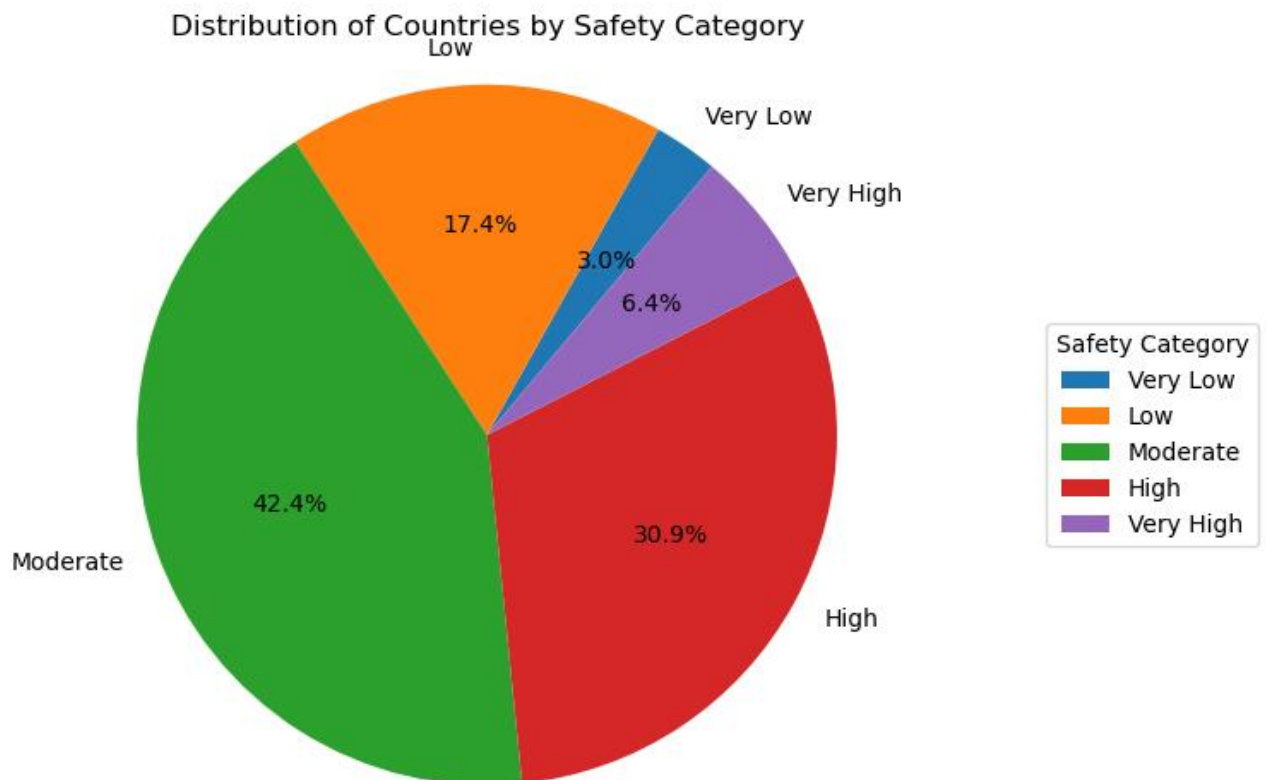


Figure 10: Pie Chart for Safety Categories.

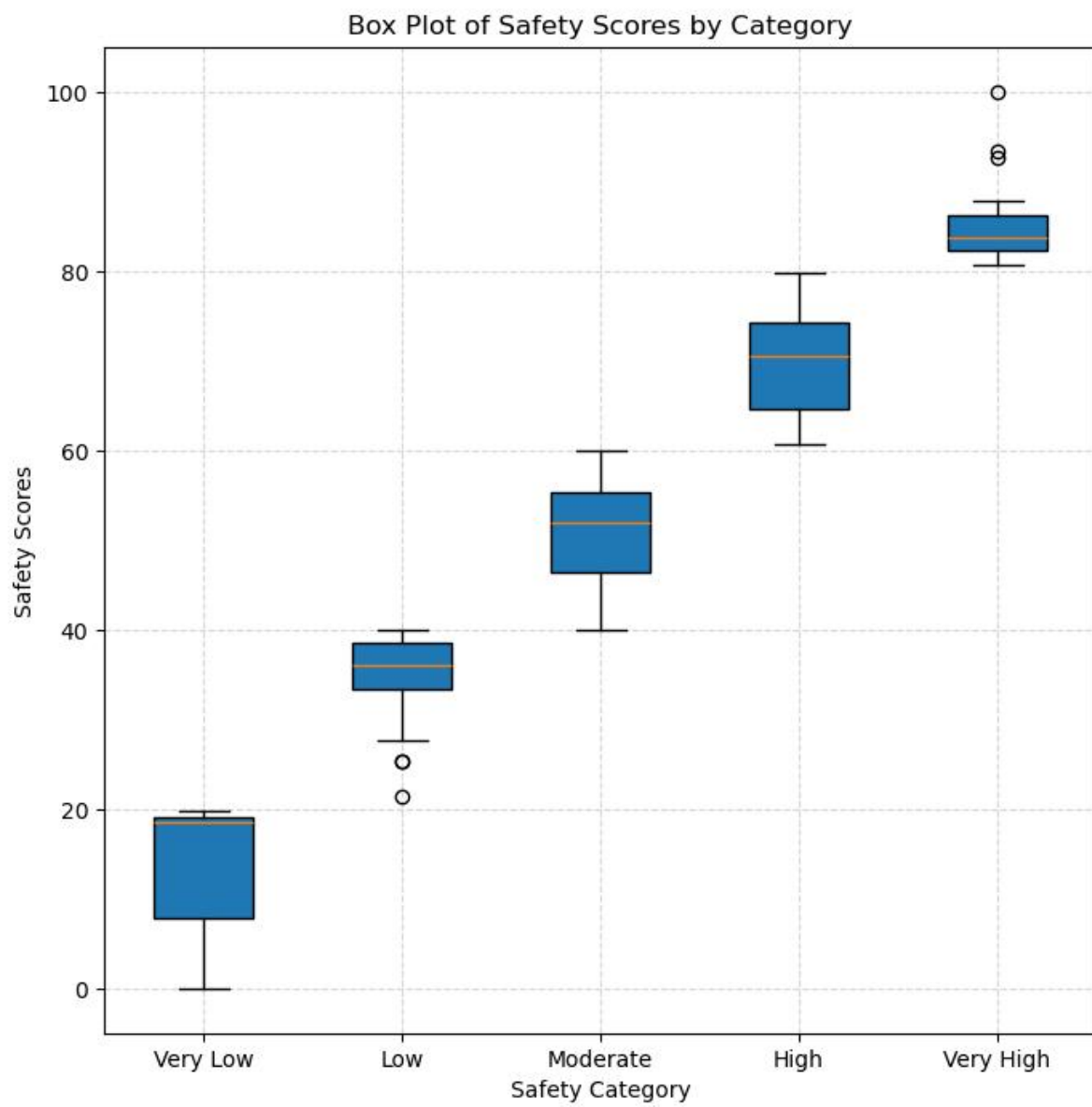


Figure 11: Box Plot for Safety Categories.

## Safety Values

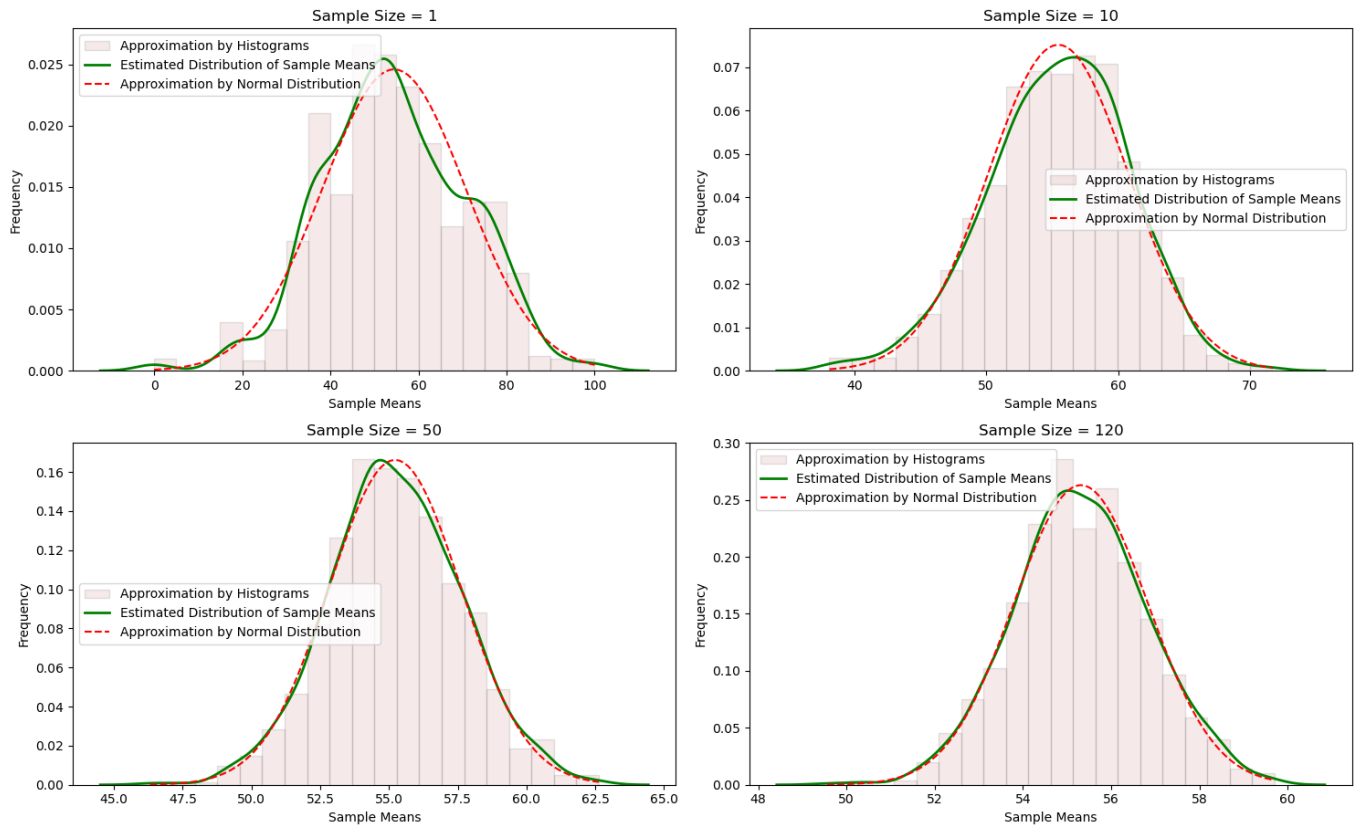


Figure 12: Approximation of Sample means for Safety values from CLT

## 4.3 For Health Care

### For Africa's Health Care Values

The average is 48.367  
The median is 47.745  
Standard Dev is 17.515  
Mean dev is 12.521  
1st Quartile is 40.655  
2nd Quartile is 47.745  
3rd Quartile is 57.328  
range is 100.000

### For Asia's Health Care Values

The average is 59.473  
The median is 61.080  
Standard Dev is 13.835  
Mean dev is 10.708  
1st Quartile is 51.812  
2nd Quartile is 61.080  
3rd Quartile is 69.910  
range is 62.260

### For Europe's Health Care Values

The average is 66.544  
The median is 67.630

Standard Dev is 11.410  
Mean dev is 9.482  
1st Quartile is 57.215  
2nd Quartile is 67.630  
3rd Quartile is 75.515  
range is 58.950

#### For North America's Health Care Values

The average is 52.721  
The median is 54.360  
Standard Dev is 19.214  
Mean dev is 13.931  
1st Quartile is 45.870  
2nd Quartile is 54.360  
3rd Quartile is 65.032  
range is 93.060

#### For South America's Health Care Values

The average is 53.174  
The median is 57.745  
Standard Dev is 19.439  
Mean dev is 14.418  
1st Quartile is 43.460  
2nd Quartile is 57.745  
3rd Quartile is 67.130  
range is 75.390

#### For Oceania's Health Care Values

The average is 39.578  
The median is 45.140  
Standard Dev is 31.485  
Mean dev is 25.753  
1st Quartile is 0.000  
2nd Quartile is 45.140  
3rd Quartile is 55.560  
range is 100.000

#### For Antarctica's Health Care Values

The average is 41.666  
The median is 38.890  
Standard Dev is 43.744  
Mean dev is 34.443  
1st Quartile is 0.000  
2nd Quartile is 38.890  
3rd Quartile is 69.440  
range is 100.000



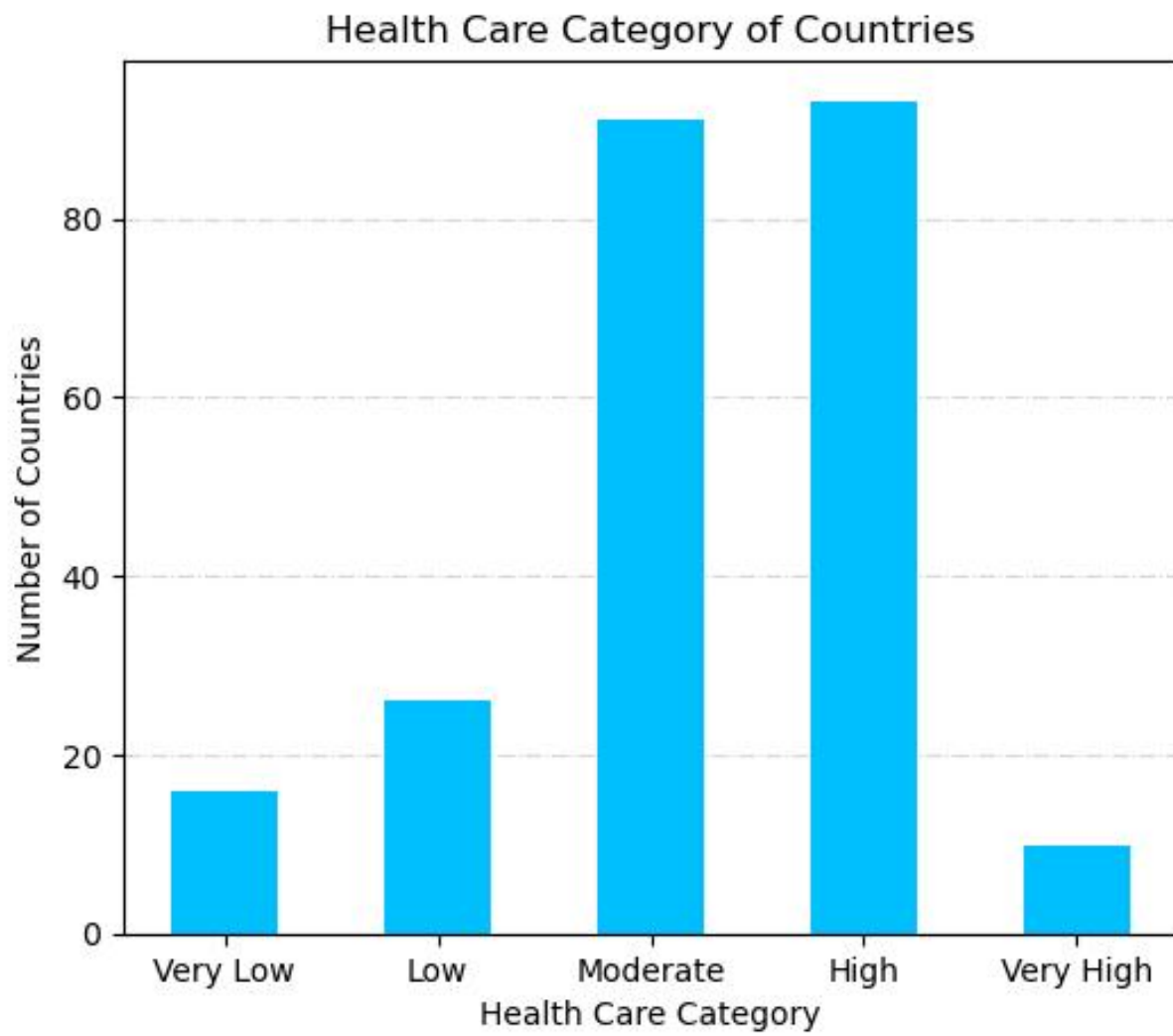


Figure 13: Bar Graph for Health Care Categories.

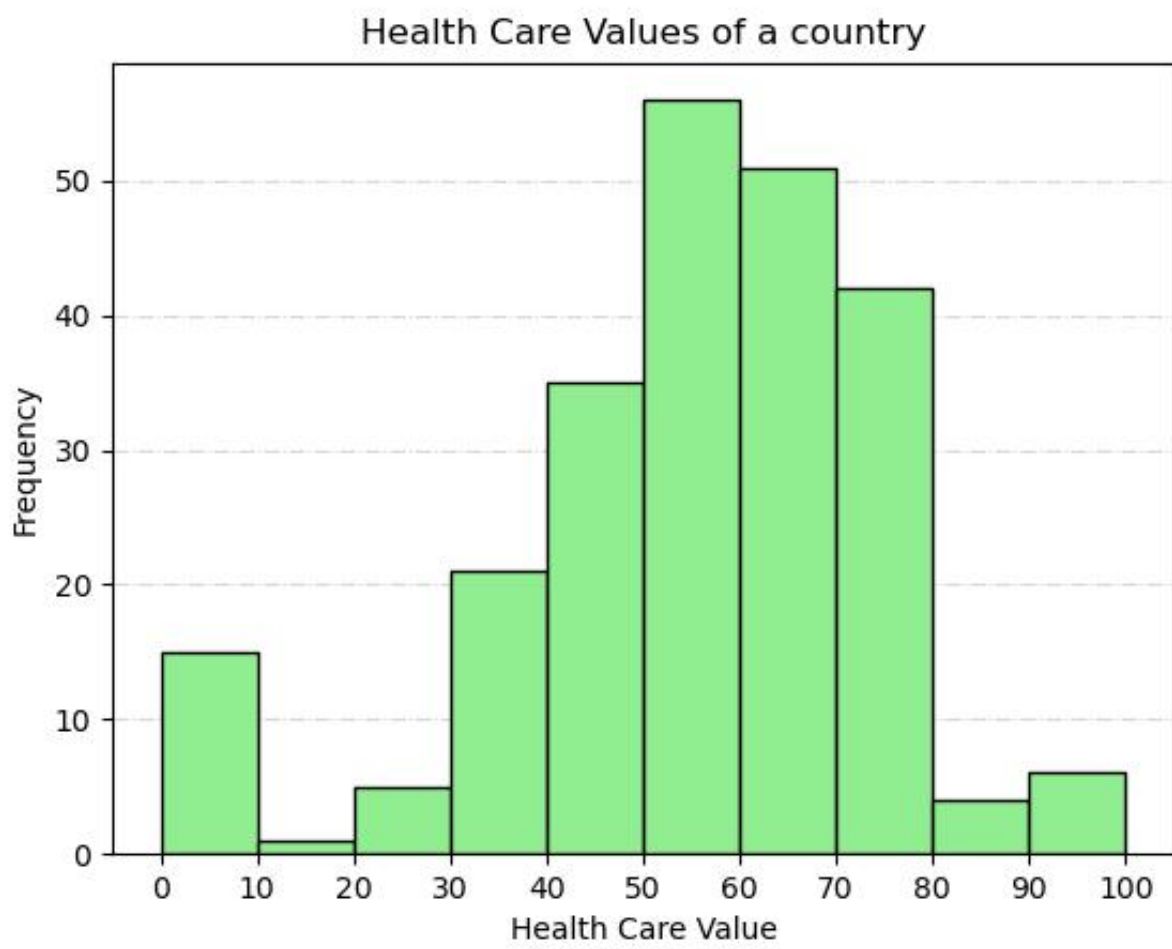


Figure 14: Histogram for Health Care Values.

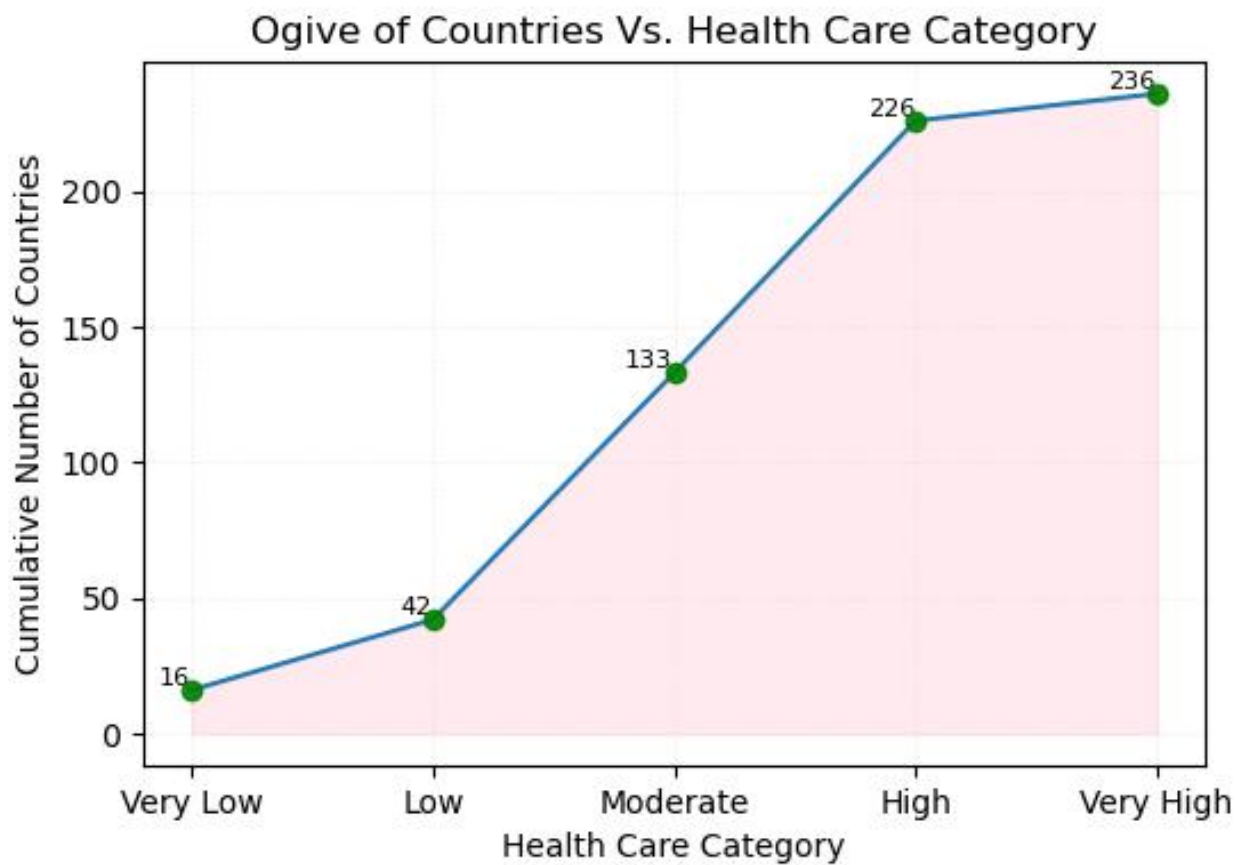


Figure 15: Ogive for Health Care Categories.

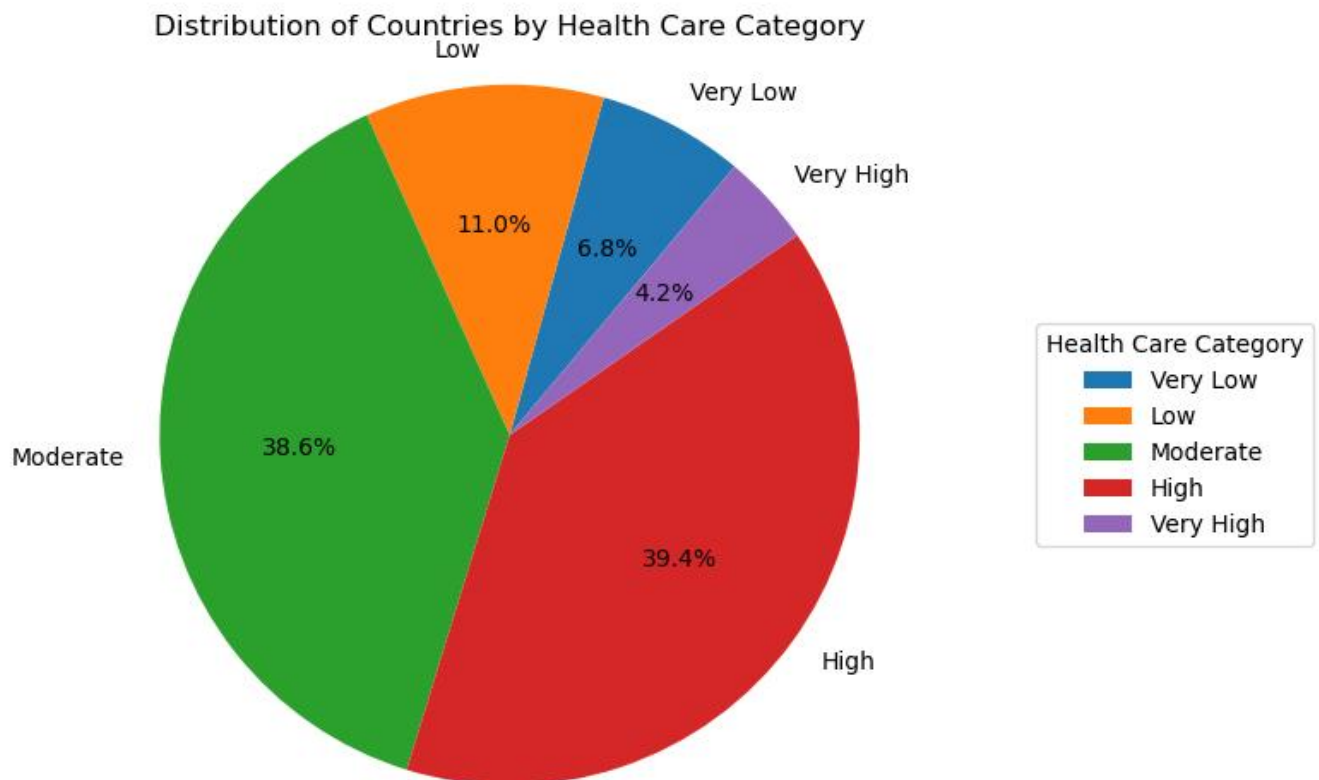


Figure 16: Pie Chart for Health Care Categories.

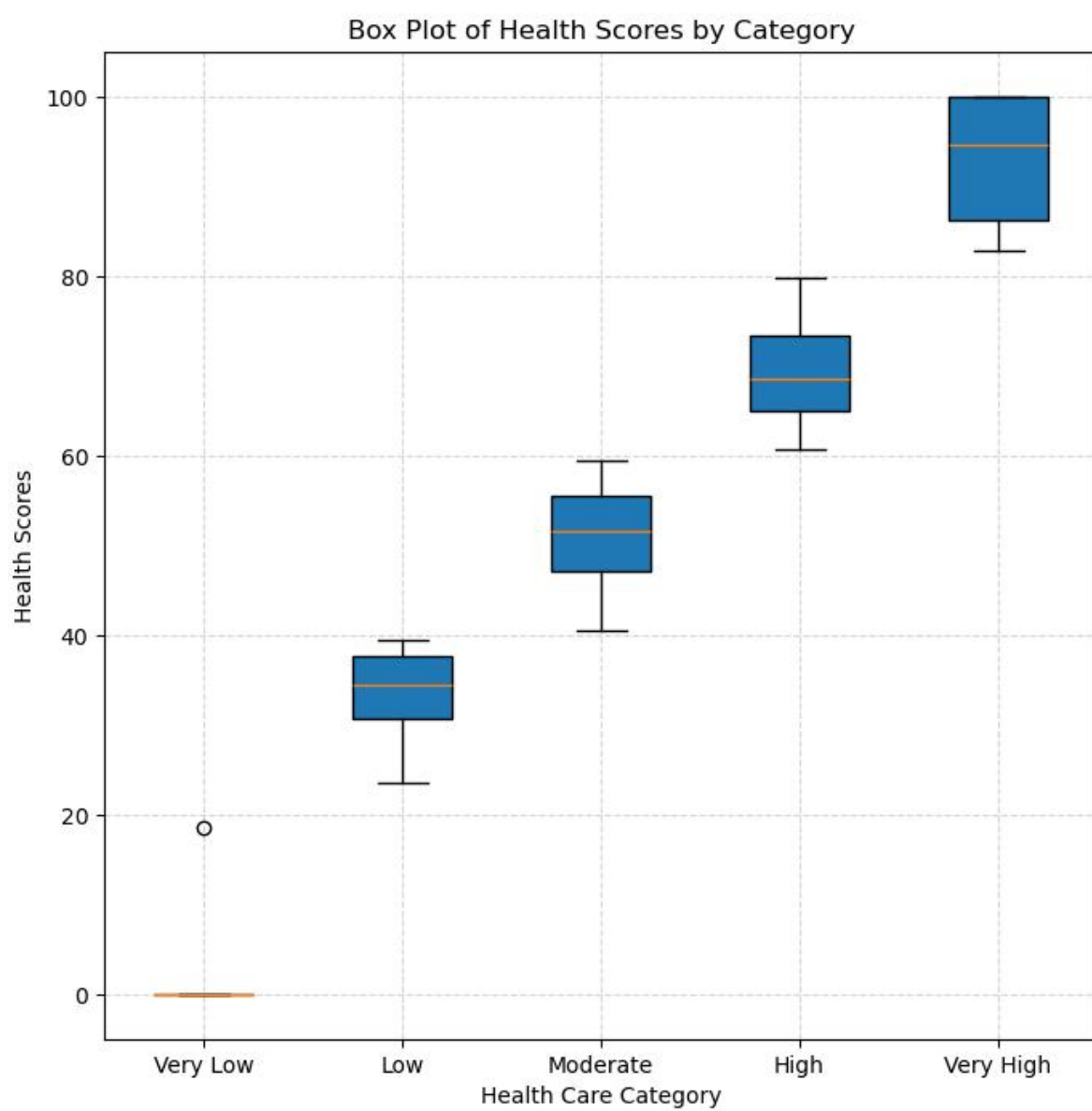


Figure 17: Box Plot for Health Care Categories.

## Health Care Values

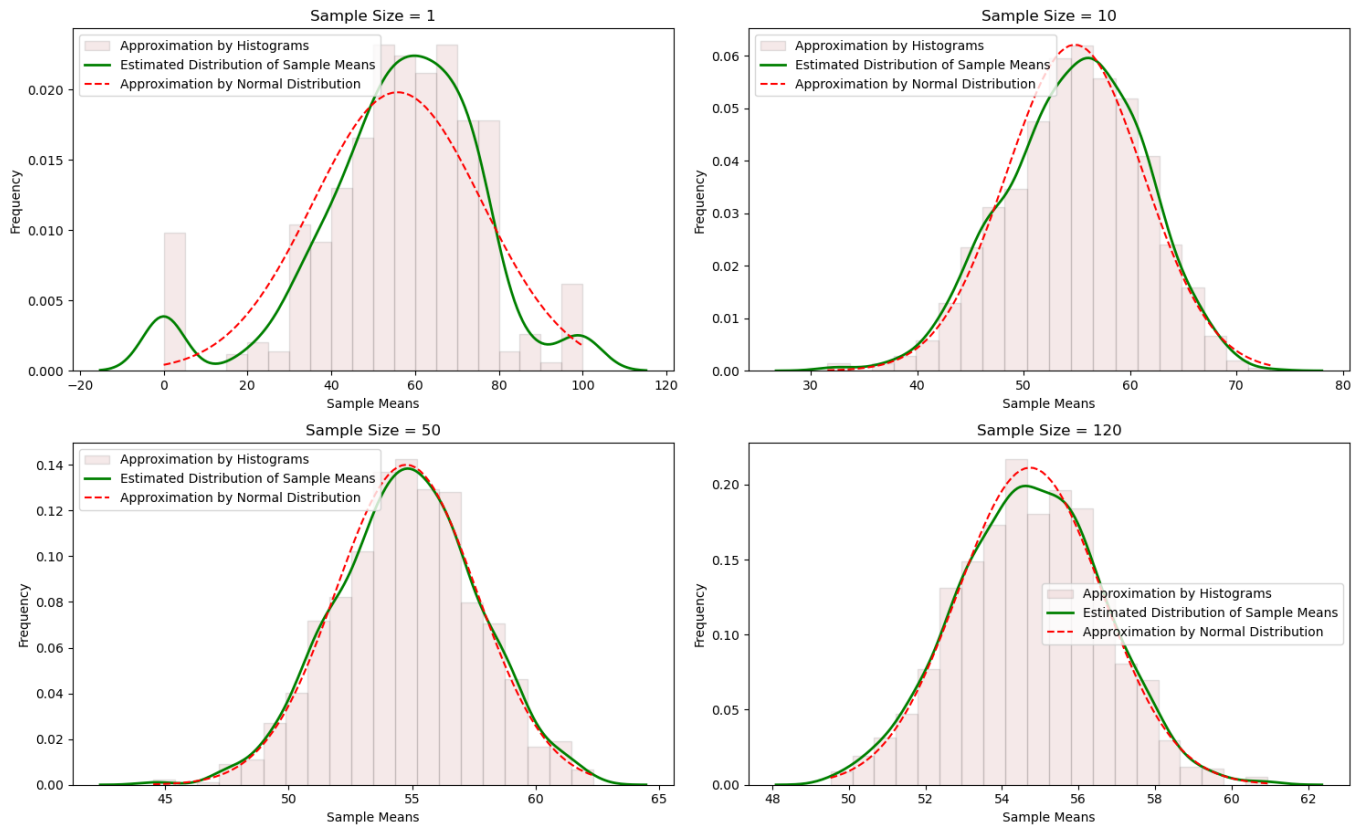


Figure 18: Approximation of Sample means for Health Care values from CLT

## 4.4 For Cost of Living (COL)

### For Africa's Cost of Living Values

The average is 48.367  
 The median is 47.745  
 Standard Dev is 17.515  
 Mean dev is 12.521  
 1st Quartile is 40.655  
 2nd Quartile is 47.745  
 3rd Quartile is 57.328  
 range is 100.000

### For Asia's Cost of Living Values

The average is 59.473  
 The median is 61.080  
 Standard Dev is 13.835  
 Mean dev is 10.708  
 1st Quartile is 51.812  
 2nd Quartile is 61.080  
 3rd Quartile is 69.910  
 range is 62.260

### For Europe's Cost of Living Values

The average is 66.544  
 The median is 67.630

Standard Dev is 11.410  
Mean dev is 9.482  
1st Quartile is 57.215  
2nd Quartile is 67.630  
3rd Quartile is 75.515  
range is 58.950

#### For North America's Cost of Living Values

The average is 52.721  
The median is 54.360  
Standard Dev is 19.214  
Mean dev is 13.931  
1st Quartile is 45.870  
2nd Quartile is 54.360  
3rd Quartile is 65.032  
range is 93.060

#### For South America's Cost of Living Values

The average is 53.174  
The median is 57.745  
Standard Dev is 19.439  
Mean dev is 14.418  
1st Quartile is 43.460  
2nd Quartile is 57.745  
3rd Quartile is 67.130  
range is 75.390

#### For Oceania's Cost of Living Values

The average is 39.578  
The median is 45.140  
Standard Dev is 31.485  
Mean dev is 25.753  
1st Quartile is 0.000  
2nd Quartile is 45.140  
3rd Quartile is 55.560  
range is 100.000

#### For Antarctica's Cost of Living Values

The average is 41.666  
The median is 38.890  
Standard Dev is 43.744  
Mean dev is 34.443  
1st Quartile is 0.000  
2nd Quartile is 38.890  
3rd Quartile is 69.440  
range is 100.000

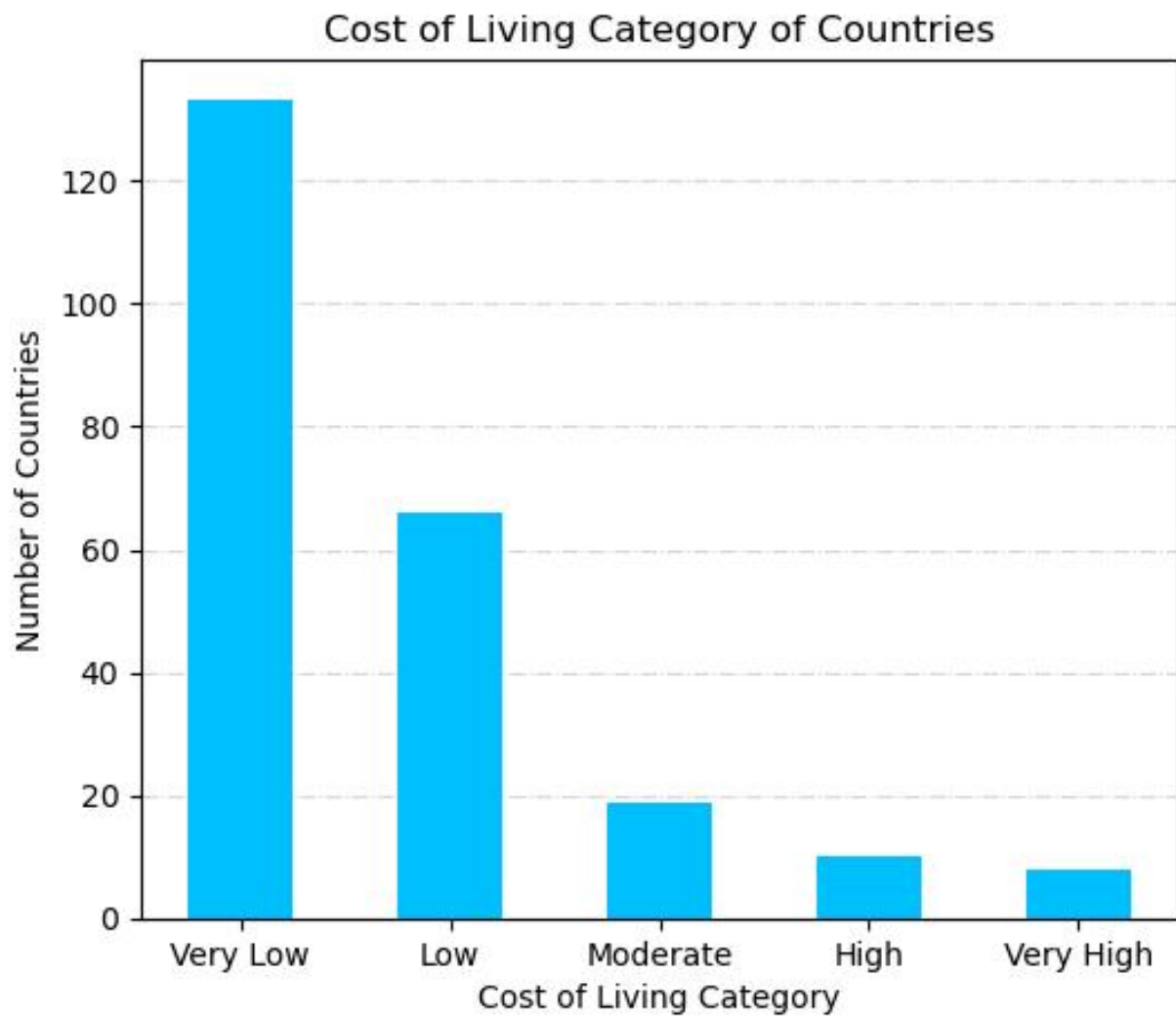


Figure 19: Bar Graph for COL Categories.

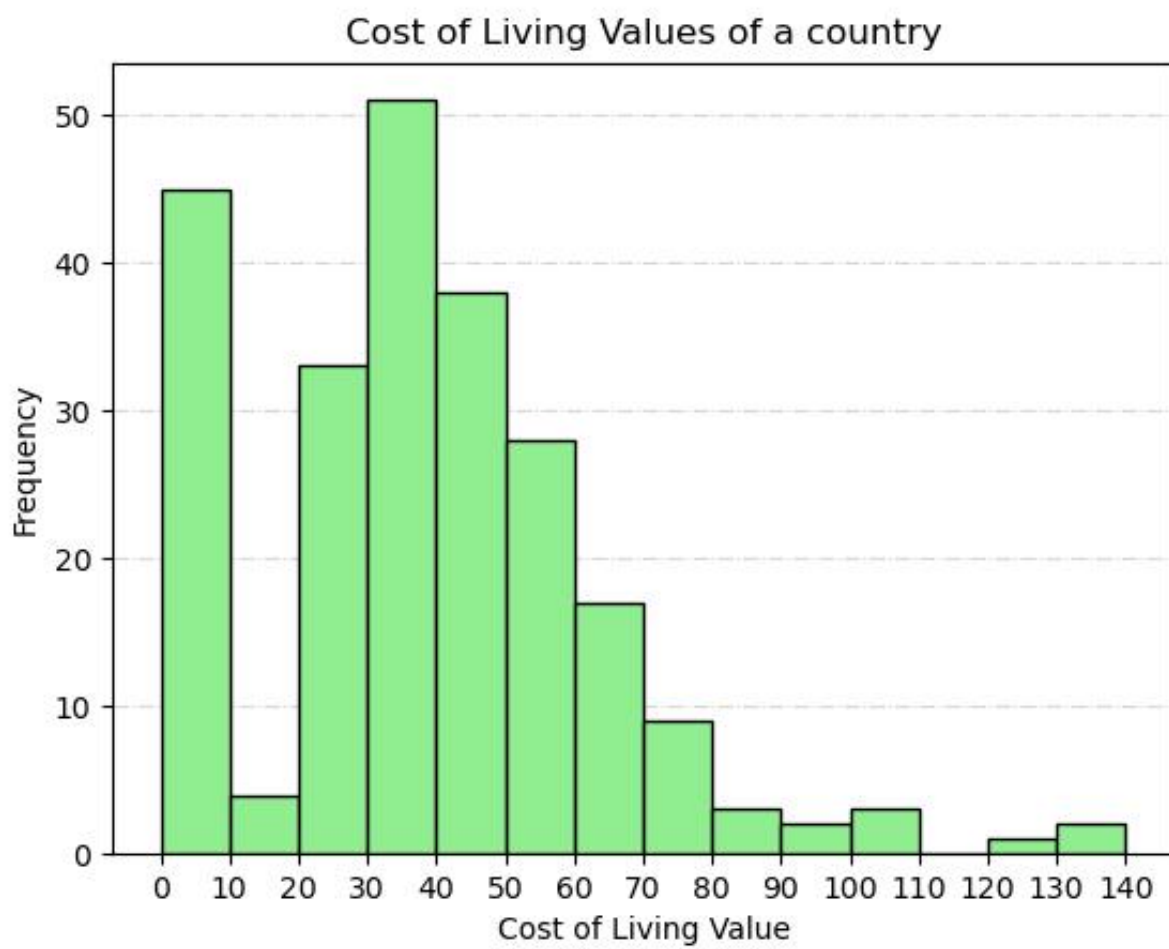


Figure 20: Histogram for COL Values.



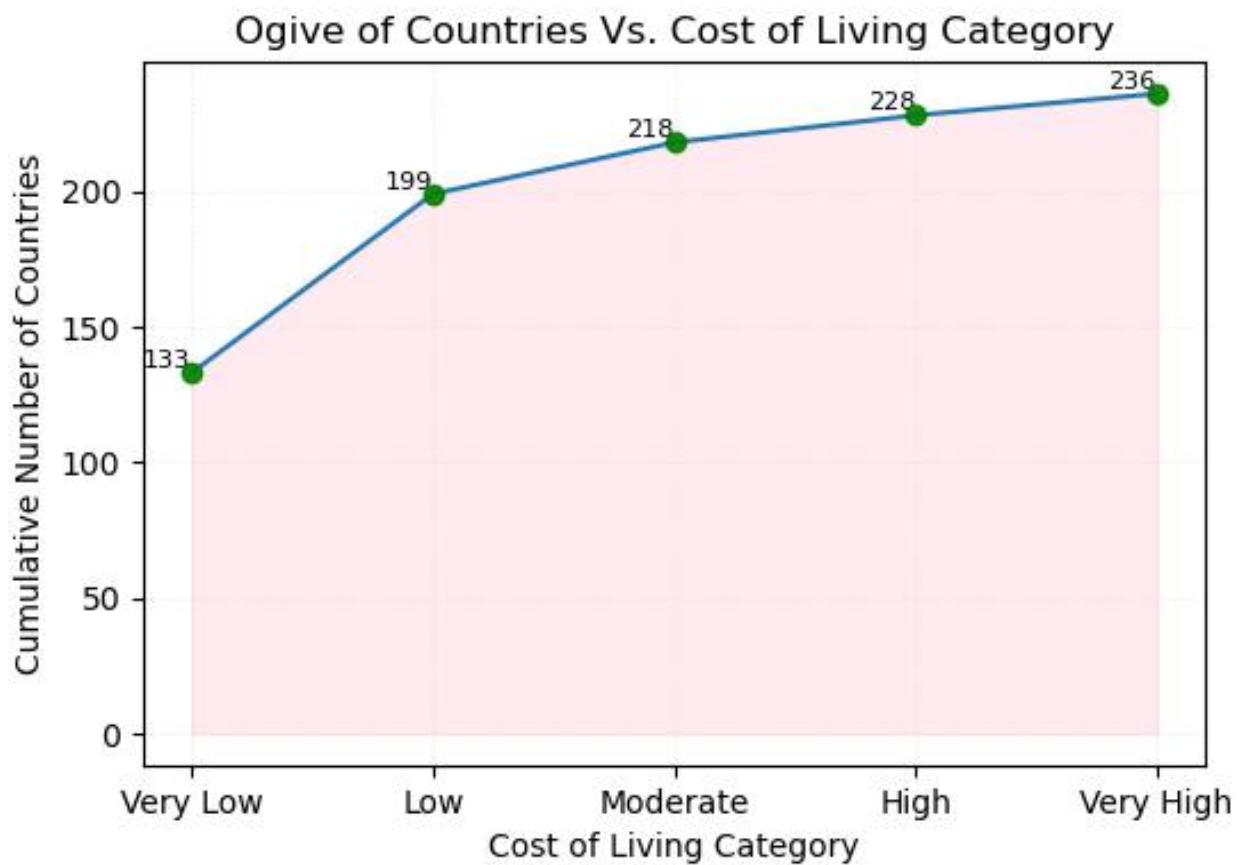


Figure 21: Ogive for COL Categories.

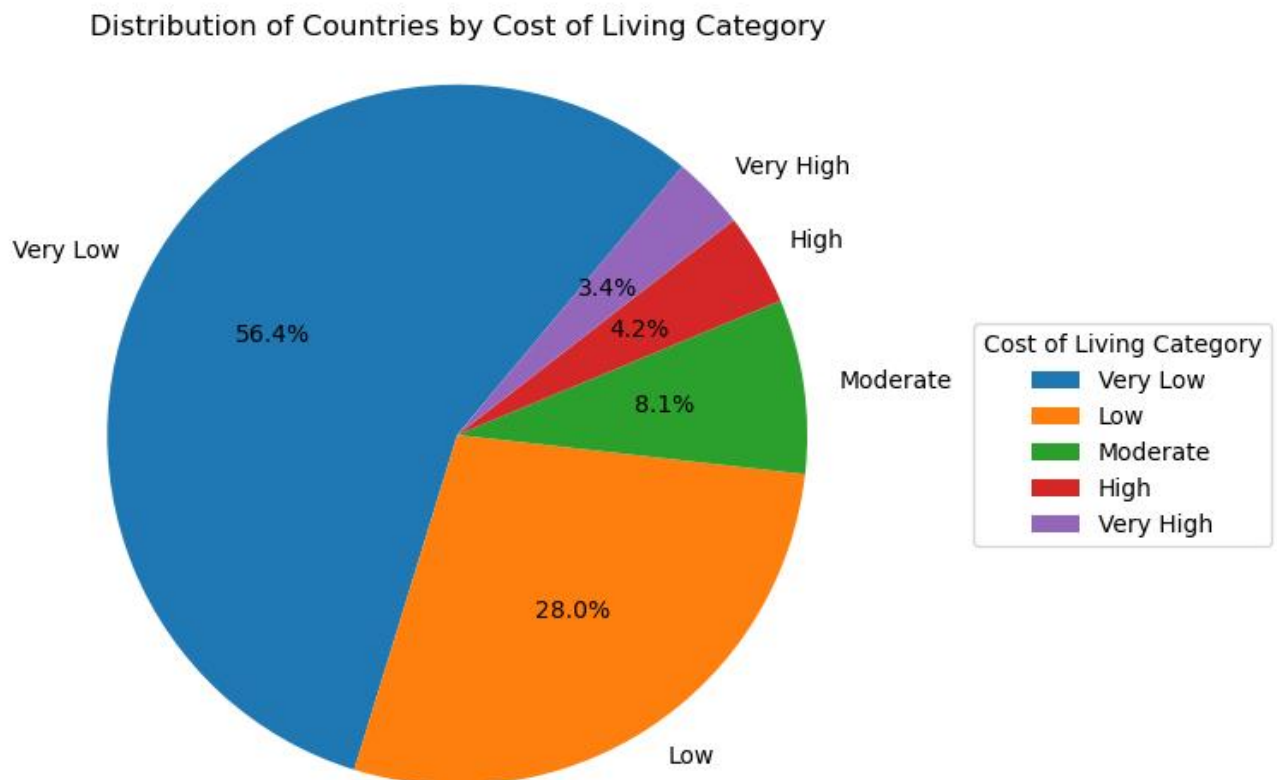


Figure 22: Pie Chart for COL Categories.

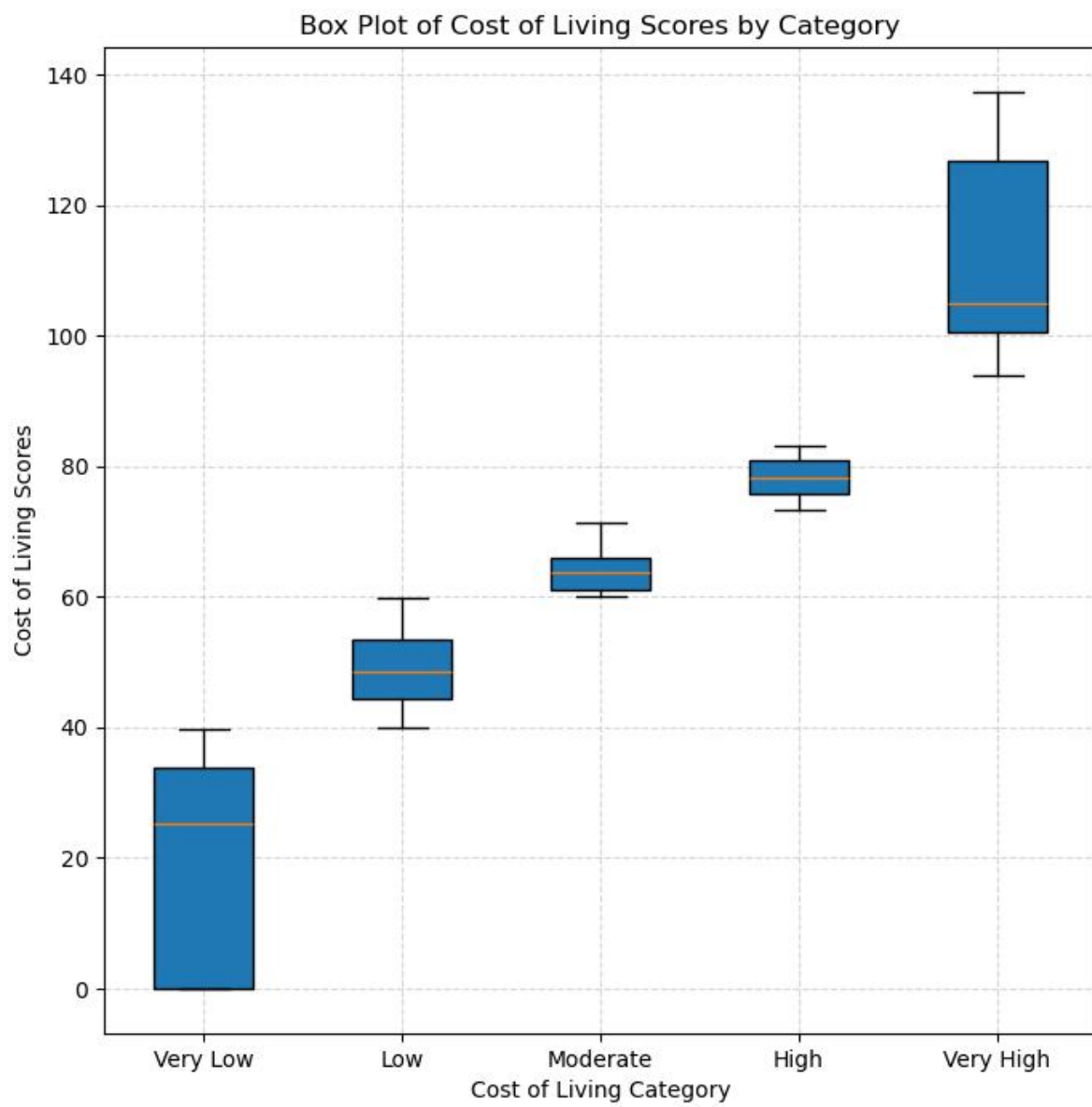


Figure 23: Box Plot for COL Categories.

## Cost of Living Values

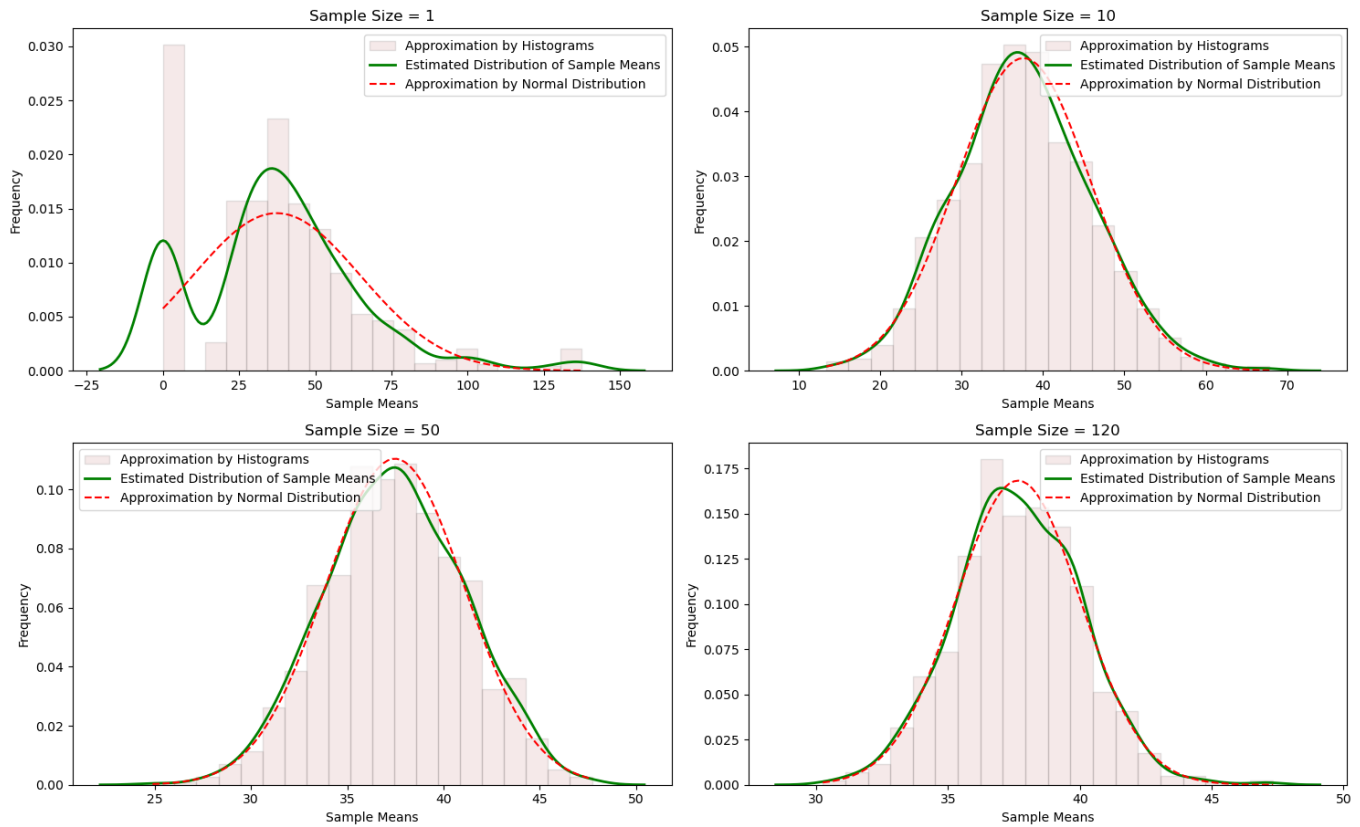


Figure 24: Approximation of Sample means for COL values from CLT

## 4.5 For Pollution

### For Africa's Pollution Values

The average is 48.367  
The median is 47.745  
Standard Dev is 17.515  
Mean dev is 12.521  
1st Quartile is 40.655  
2nd Quartile is 47.745  
3rd Quartile is 57.328  
range is 100.000

### For Asia's Pollution Values

The average is 59.473  
The median is 61.080  
Standard Dev is 13.835  
Mean dev is 10.708  
1st Quartile is 51.812  
2nd Quartile is 61.080  
3rd Quartile is 69.910  
range is 62.260

### For Europe's Pollution Values

The average is 66.544  
The median is 67.630

Standard Dev is 11.410  
Mean dev is 9.482  
1st Quartile is 57.215  
2nd Quartile is 67.630  
3rd Quartile is 75.515  
range is 58.950

#### For North America's Pollution Values

The average is 52.721  
The median is 54.360  
Standard Dev is 19.214  
Mean dev is 13.931  
1st Quartile is 45.870  
2nd Quartile is 54.360  
3rd Quartile is 65.032  
range is 93.060

#### For South America's Pollution Values

The average is 53.174  
The median is 57.745  
Standard Dev is 19.439  
Mean dev is 14.418  
1st Quartile is 43.460  
2nd Quartile is 57.745  
3rd Quartile is 67.130  
range is 75.390

#### For Oceania's Pollution Values

The average is 39.578  
The median is 45.140  
Standard Dev is 31.485  
Mean dev is 25.753  
1st Quartile is 0.000  
2nd Quartile is 45.140  
3rd Quartile is 55.560  
range is 100.000

#### For Antarctica's Pollution Values

The average is 41.666  
The median is 38.890  
Standard Dev is 43.744  
Mean dev is 34.443  
1st Quartile is 0.000  
2nd Quartile is 38.890  
3rd Quartile is 69.440  
range is 100.000

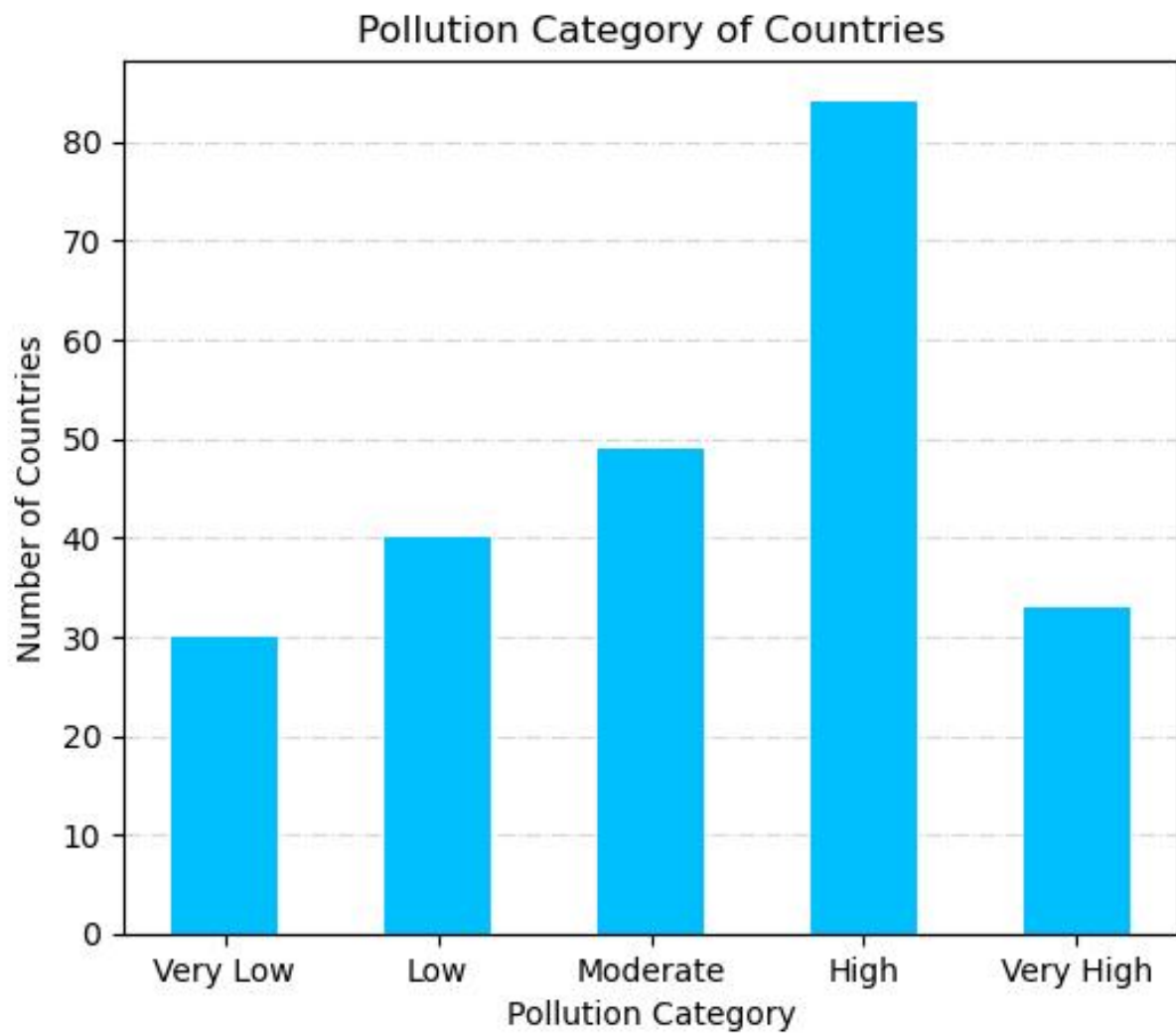


Figure 25: Bar Graph for Pollution Categories.

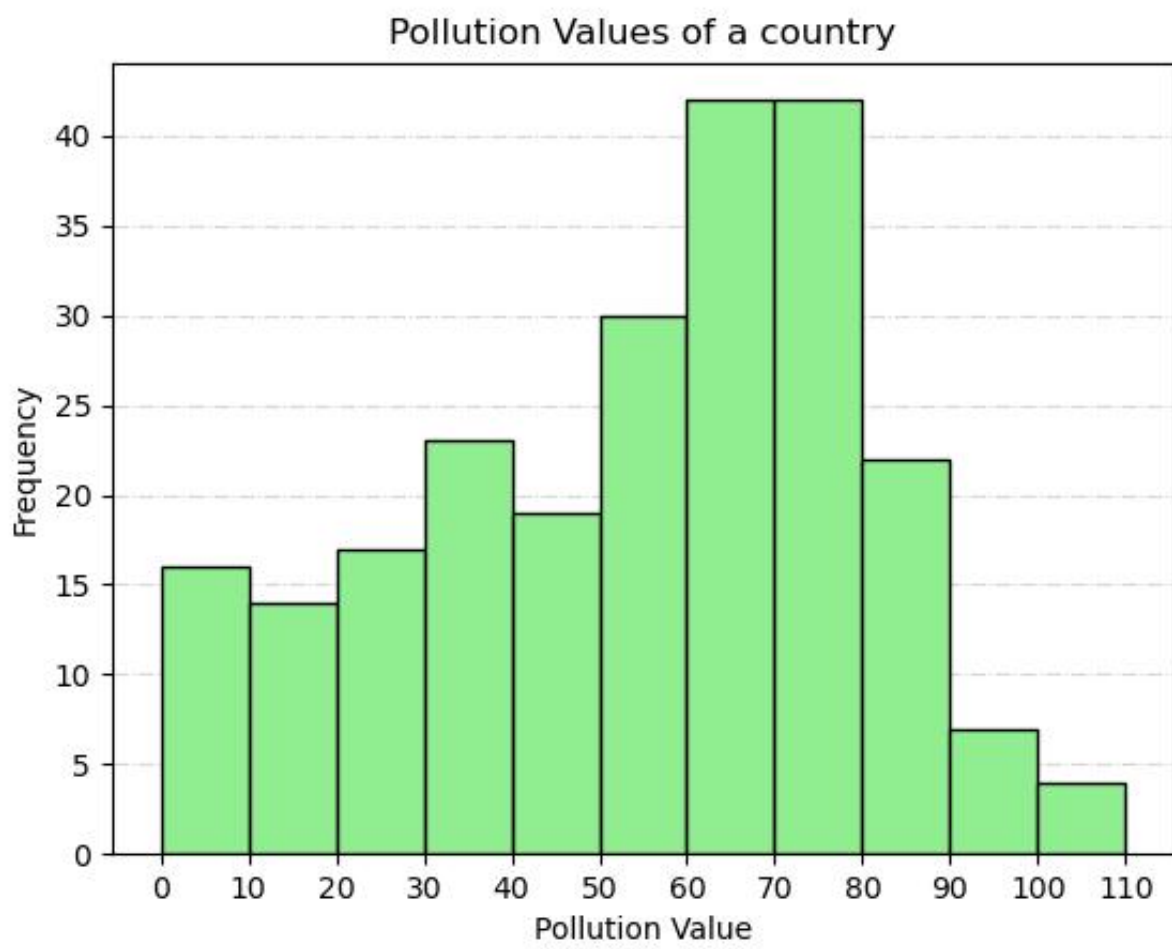


Figure 26: Histogram for Pollution Values.

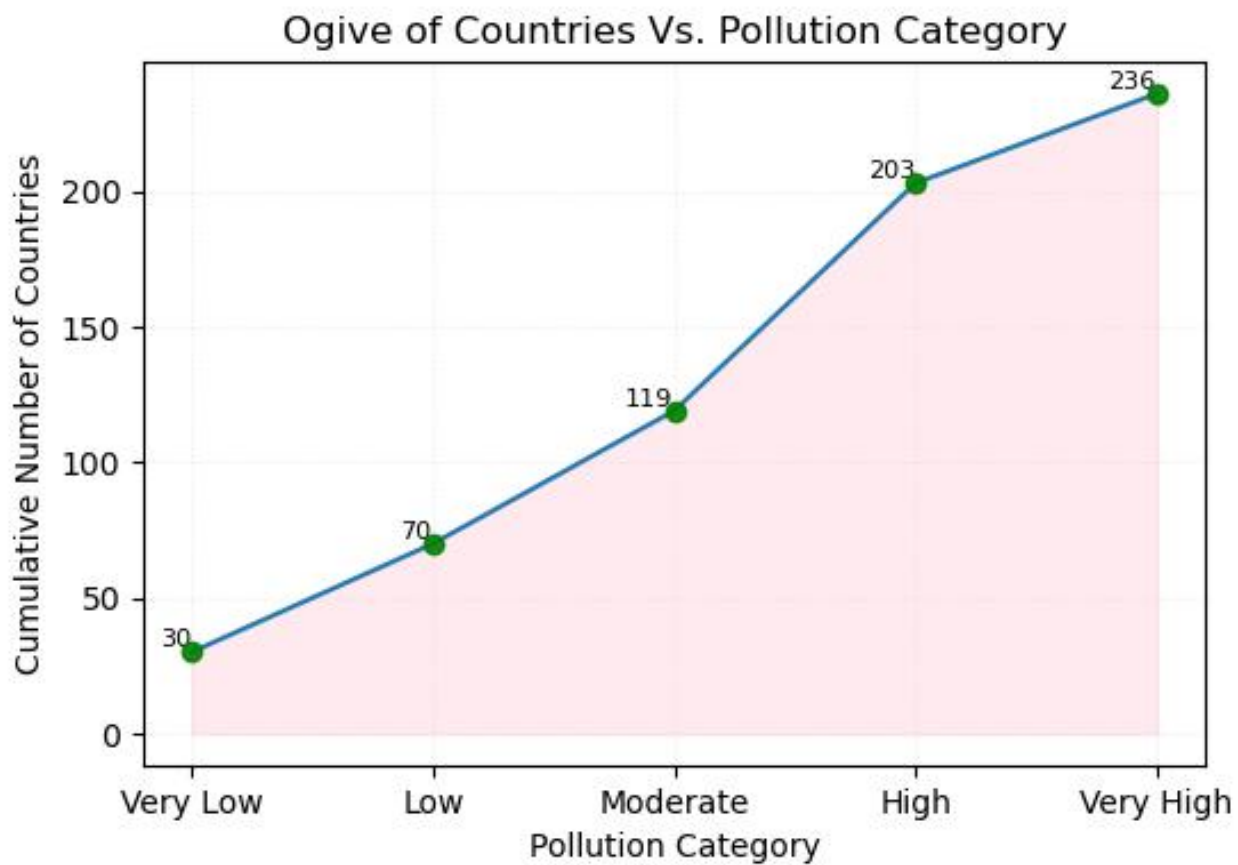


Figure 27: Ogive for Pollution Categories.

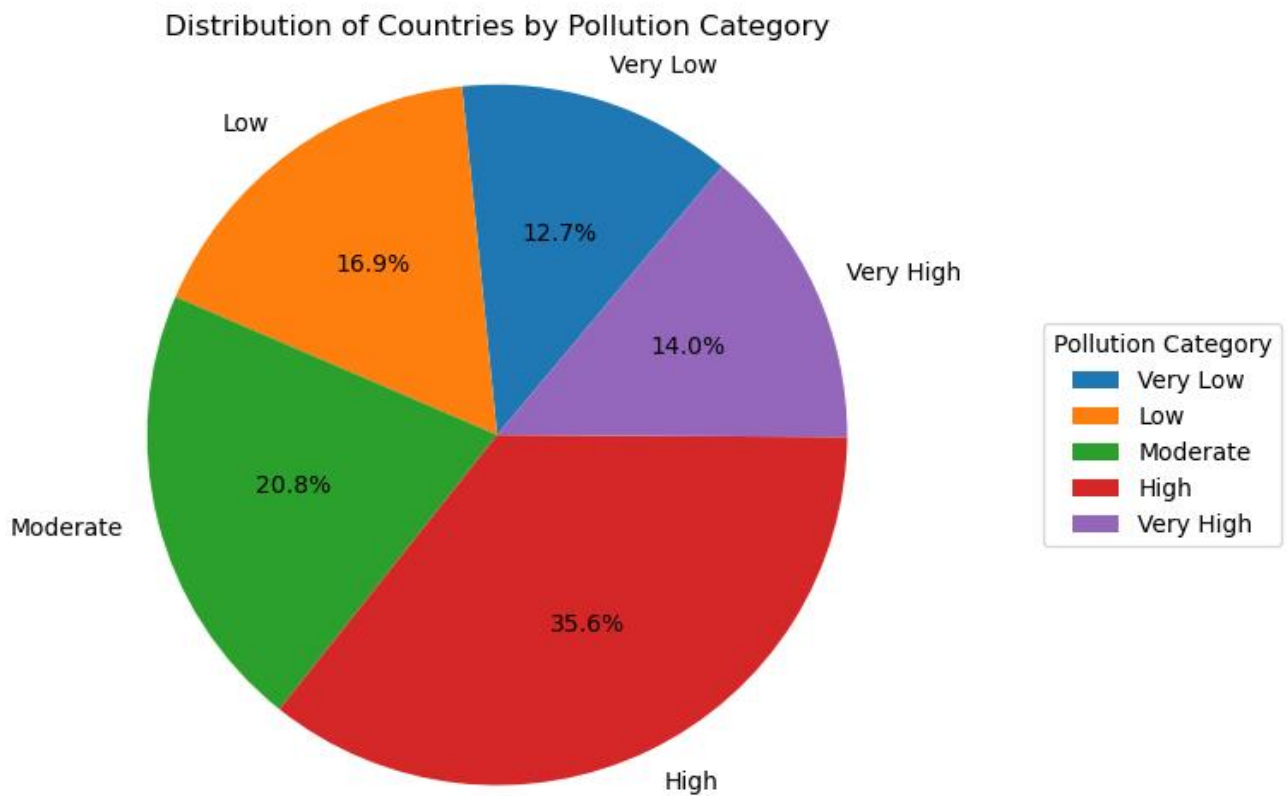


Figure 28: Pie Chart for Pollution Categories.

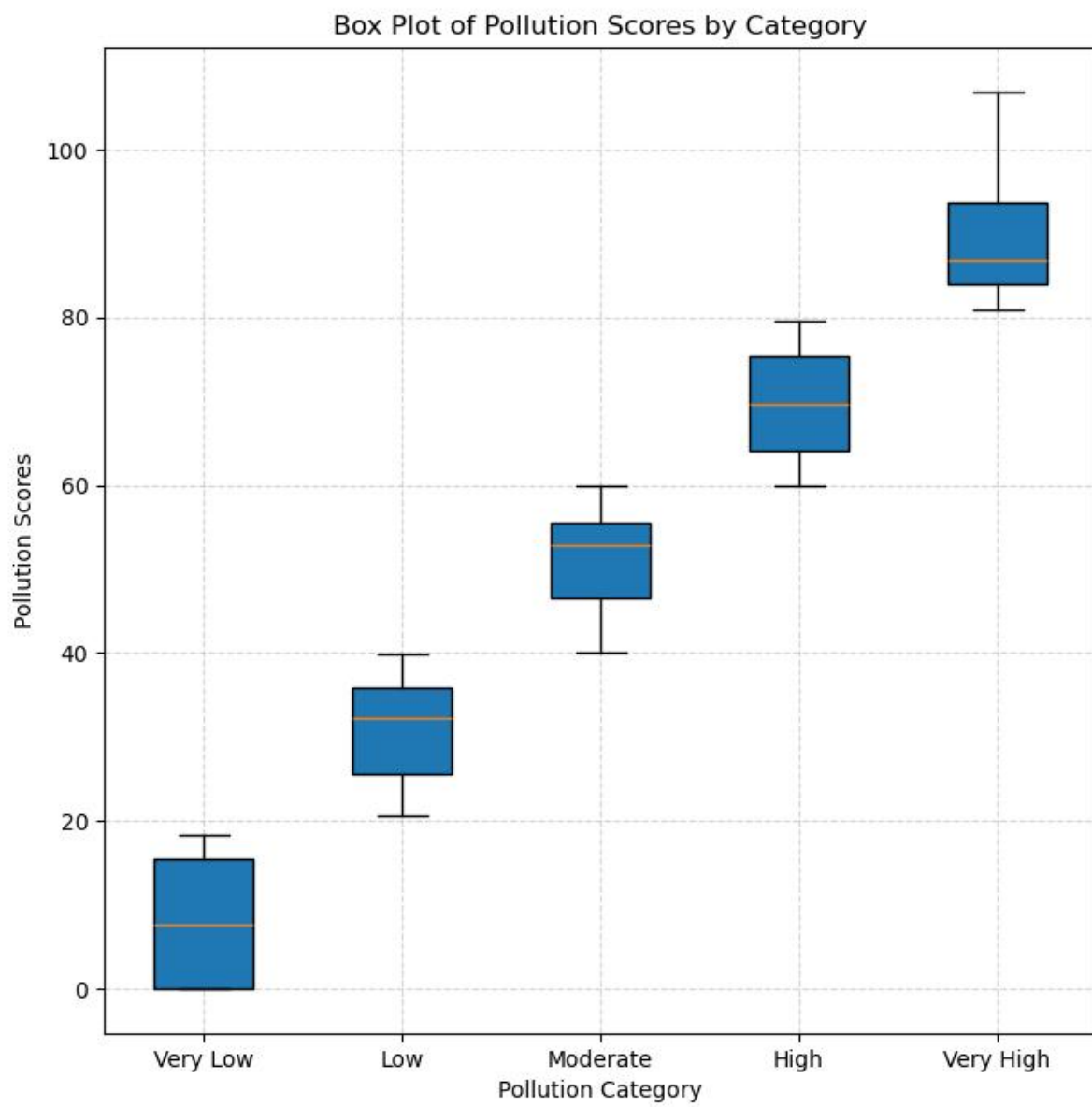


Figure 29: Box Plot for Pollution Categories.



## Pollution Values

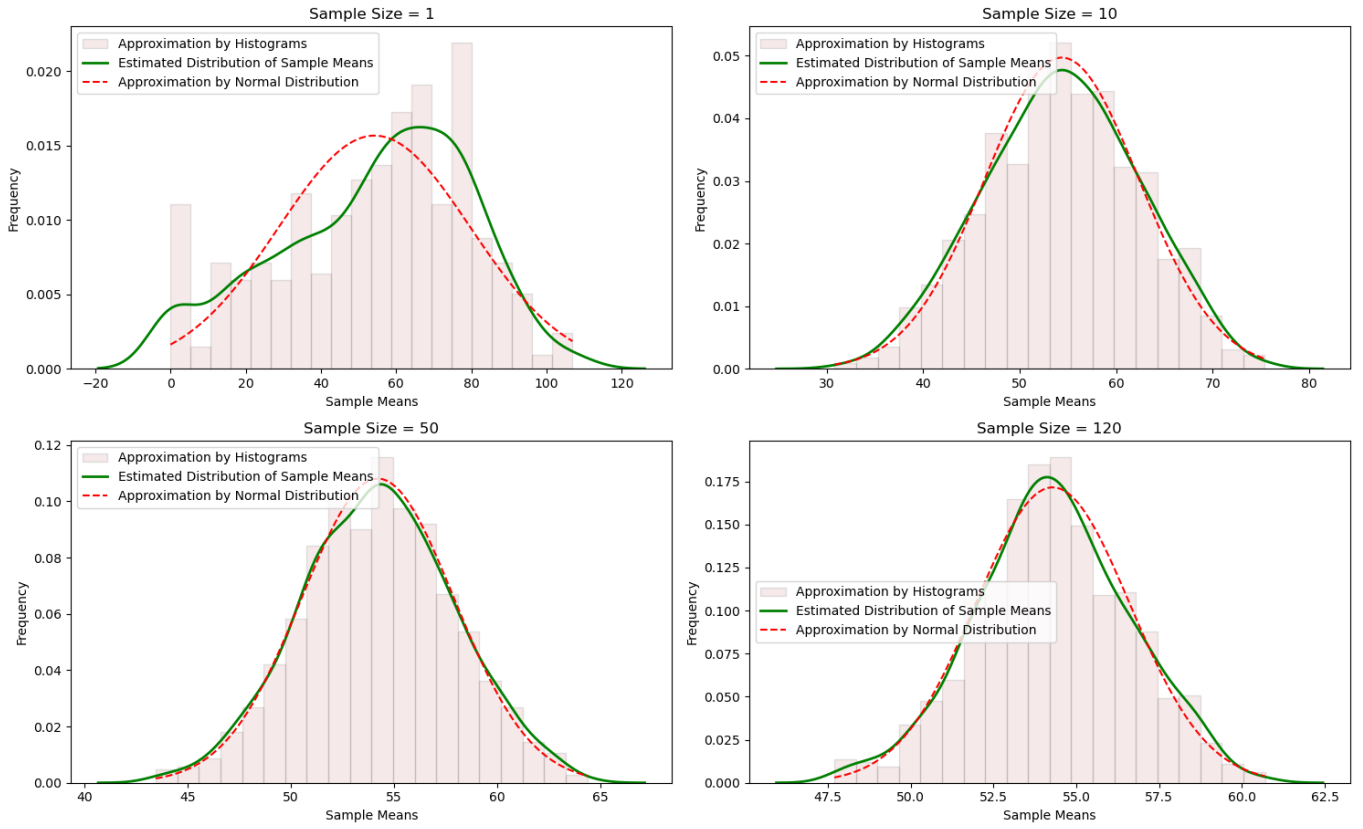


Figure 30: Approximation of Sample means for Pollution values from CLT

## 4.6 Co-relation Matrix between features

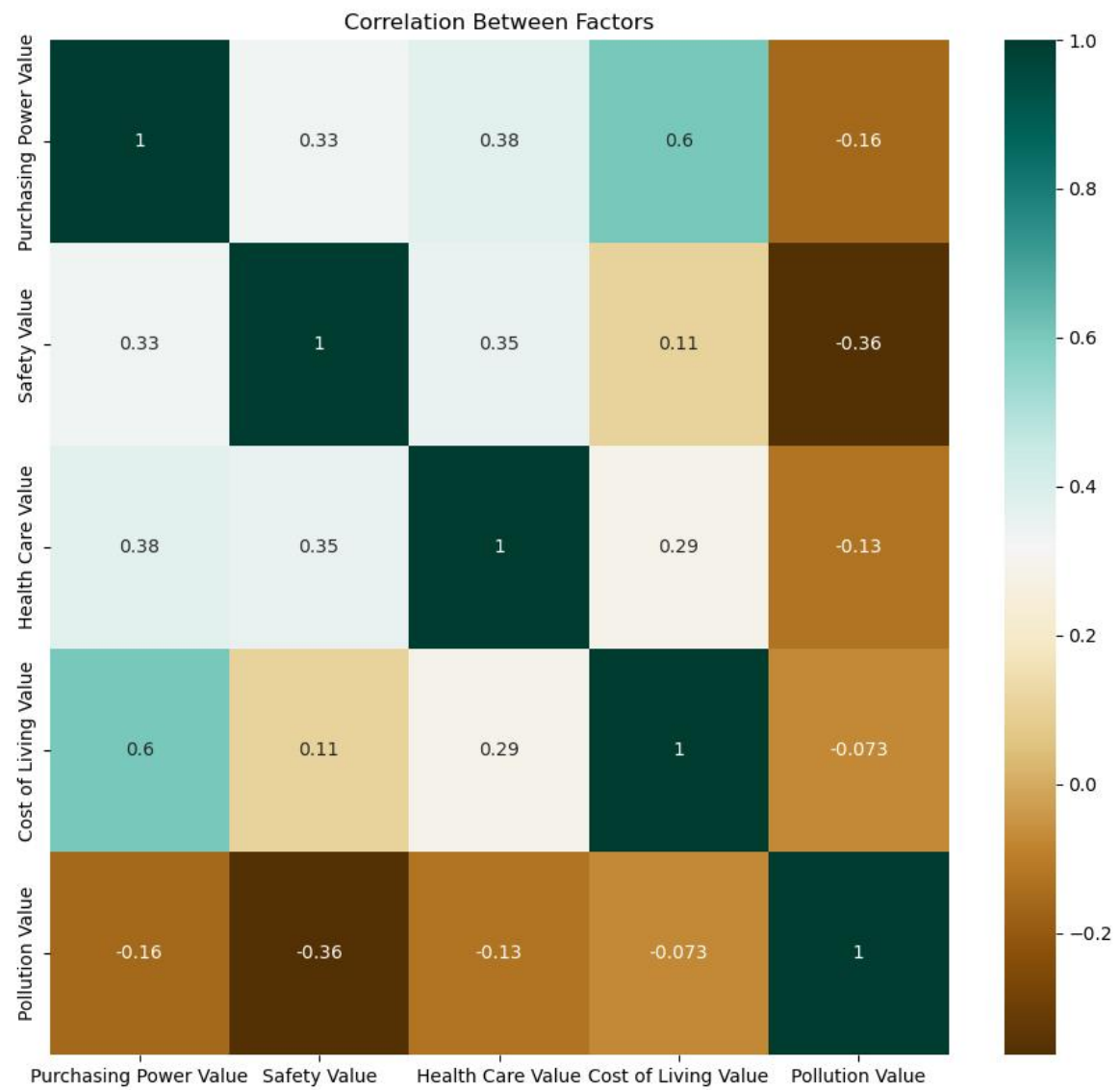


Figure 31: Correlation Between Factors

## **5 Observations from Visualizations**

### **5.1 Purchasing Power Distribution**

Most countries have low to moderate purchasing power. Some outliers have significantly high purchasing power.

### **5.2 Safety Index Distribution**

The safety values follow a more normal distribution, with most countries around the 50-70 range. Some extreme values indicate very unsafe or very safe countries.

### **5.3 Health Care Index Distribution**

Many countries have healthcare scores clustered around 50-70. Some countries have exceptionally high healthcare scores (close to 100).

### **5.4 Purchasing Power vs. Cost of Living**

There is a positive correlation: Countries with a higher purchasing power tend to have a higher cost of living. However, some countries have a high cost of living but relatively low purchasing power, indicating economic imbalances.

### **5.5 Pollution Index Distribution**

Pollution scores are mostly spread between 20-80. Some countries have extremely high or low levels of pollution.

### **5.6 Correlation Insights (from Heatmap)**

Purchasing Power and Cost of Living have a strong positive correlation. Safety and Health Care are moderately correlated, meaning that safer countries often have better healthcare. Pollution is weakly correlated with other factors, meaning that a country's economy does not always predict its pollution level.