**Evaluation of Public Knowledge on the Impact of Climate Change on Emerging Infectious Diseases.**

Author: Elohor Okpako.

Table of Contents:

1. Introduction
2. Data preprocessing
3. Data Exploration
4. Data Exploration and Statistical Tests
5. Conclusion

**Introduction**:

The intertwining complexities of climate change and emerging infectious diseases have become increasingly evident in recent years, prompting a critical need to assess public understanding of their interconnectedness. As the global climate continues to undergo significant shifts, the frequency and distribution of infectious diseases are also evolving, presenting profound challenges to public health. This study seeks to evaluate the level of public knowledge regarding the impacts of climate change on the emergence and spread of infectious diseases. By exploring public awareness, perceptions, and misconceptions surrounding this crucial intersection, we aim to contribute to literature and inform targeted educational initiatives and policy interventions geared towards mitigating the adverse health outcomes associated with climate-driven infectious disease dynamics.

**Data Preprocessing**

**Data Collection:**

The data utilized in this study were gathered through online questionnaires distributed via Google Forms across various social media platforms, including WhatsApp, Facebook, and LinkedIn. A total of 231 unique responses were collected from participants residing in 28 different countries spanning four continents.

**Data Quality Assurance:**

To ensure representativeness, different countries were sampled in the study. Additionally, the collected data underwent a rigorous process of data wrangling to address inconsistencies and prepare it for analysis. Notably, this involved standardizing responses, particularly for open-ended questions such as specifying country of origin and education level.

**Scoring Methodology:**

Each participant's responses were scored using a standardized criterion, whereby correct answers were awarded a score of +1, while incorrect choices incurred a penalty of -0.5. This scoring approach was implemented to mitigate the impact of random guessing bias. No deductions were made for responses indicating uncertainty or lack of knowledge. For further details on the scoring guide, refer to the provided sample on the project's GitHub page.

**Data Exploration**

The Respondents from the survey were younger adults as represented in the histogram below with mean age of 28:

A graph with blue squares

Description automatically generated

Most of the respondents from the survey were from Africa with Nigeria having the highest participants (70), this is followed by Uganda (61), Ethiopia (38), Eritrea (27) and others.

A map of the world

Description automatically generated

The score distribution of the participants was a uniform distribution as shown below:

A graph with blue squares

Description automatically generated

The gender distribution of the participants is shown below with 60% male and 40% female.

A graph of a bar chart

Description automatically generated with medium confidence

**Data Exploration and Statistical Tests**

To delve deeper into the research inquiry and gain insight into participant responses, the datasets were scrutinized with the following inquiries:

### Q1. Is the Score performance a function of age?

### A screen shot of a screen Description automatically generated

### The scatter plot above shows that there is no correlation between the participants age and the scores they obtained.

### Q2: What is the mean performance by Gender?

### The chart below shows that there is no significant difference in the performance by Gender. The average scores by Male and Female was approximately 60%.

A screen shot of a bar chart

Description automatically generated

A screenshot of a computer

Description automatically generated

With a p-value > .05, we further confirm that there is no significant difference between the scores for the male and female participants.

### Q3 What is the mean score performance by Field of study? (Life Science, Non Life Science)

A screen shot of a graph

Description automatically generated

We observe some differences in the mean score between Life Sciences and Non-Life Sciences as shown with the boxplot.

A screenshot of a computer code

Description automatically generated  
P\_value<0.05, there is a significant difference between the knowledge of life science students and non\_life science students,

### Q4 Is there a difference in the performance by Level of study? (Bachelors, Post-graduates, Others)?

A screen shot of a computer screen

Description automatically generated

We observe that the mean score by Level of Study appears not to be different.

### Q5 What is the mean knowledge score by Continents, and top represented countries?

A screen shot of a graph

Description automatically generated

We observe that Africa had more respondents, this is followed by Europe. However, from the boxplot, we observe that Africa performed less as compared to other continents. However, there may be a bias due to the sample size of the representatives from each of the continents.

However, since we had more participants from Africa, I selected 4 of the top represented countries from Africa to find out if there was a difference in the knowledge of the incidence of climate change on EID among them.

The results are shown below:

A screen shot of a graph

Description automatically generated

To further understand if there was a statistical difference among the countries, a Tukey test was performed.

A screenshot of a computer

Description automatically generated

Here, we observe that there was a difference when each of the countries was compared with Uganda, except Nigeria showing that there the samples observed from Eritrea and Ethiopia possess more knowledge about the impact of Climate change on EID compared to Nigeria and Uganda.

### Q6: What percentage of the participants are knowledgeable and not knowledgeable?

A screenshot of a graph

Description automatically generated

Only about 14% of the participants were knowledgeable about the questions which were used for assessment on this study and setting a scoring threshold of 80%.

### Q7: What is the distribution of the participants knowledge on some of the infectious diseases?

A graph with blue bars

Description automatically generated

According to the questions which were asked, almost all the participants had heard about COVID, and Ebola and very few had knowledge about Nosocomial infections.

**Q8: Which information source did the participants get information about the effect of climate change on zoonosis?**

A bar graph with blue and black text

Description automatically generated

The study also shows that most of the participants obtained information from the Internet and social media to obtain information on the incidence of climate change and emerging infectious diseases.

**Conclusion**

In conclusion, our findings highlight a notable disparity in knowledge levels regarding the impact of climate change on the emergence of infectious diseases between participants from non-life science backgrounds compared to those from life science backgrounds. Interestingly, our analysis did not reveal any significant differences in knowledge levels across various educational levels. However, a substantial variance was observed among the most represented countries, suggesting potential influences stemming from sociodemographic backgrounds and regions identified as emerging infectious disease (EID) hotspots. Overall, our study underscores that the general public remains inadequately informed about the intricate interplay between climate change and the emergence of infectious diseases.