**DANA 4820 TEAM PROJECT – PROPOSAL**

**Group 6 Members:**

|  |  |  |  |
| --- | --- | --- | --- |
| Hassan Zhang | Dawn Dang | Mya Nandar May | Pinky  Sa-ngarmangkang |
| 100411767 | 100413772 | 100282021 | 100416165 |

**Population:** The population is made up of individuals whose carbon footprints are being measured or estimated.

**Sample size:** 10,000 individuals whose carbon footprints are being measured or estimated.

**Objective:** To perform statistical analysis in R using chi-squared, odds ratio in a three-way table, and generalized linear model to answer the questions of interest below.

**Questions:**

1. Do dietary choices have a notable association with carbon emission levels? How does "Social Activity" impact the relationship between diet and carbon emission levels?
2. What is the relationship between the number of clothes purchased per month and the monthly groceries bill when analyzed using a generalized linear model?
3. Does a linear relationship exist between levels of social activity and individuals' decisions to utilize energy efficiently?

**Variables:**

|  |  |  |
| --- | --- | --- |
| **Variable Names** | **Description** | **Variable Types** |
| Diet | Diet | Categorical |
| Social Activity | Frequency of participating in social activities | Categorical |
| Energy Efficiency | Whether or not you care about purchasing energy efficient devices | Categorical |
| Carbon Emission | Total carbon emissions | Numerical |
| How Many New Clothes Monthly | Number of clothes purchased monthly | Numerical |
| Monthly Grocery Bill | Monthly amount spent on groceries, in dollars | Numerical |

**Source of dataset: KAGGLE**

[**https://www.kaggle.com/datasets/dumanmesut/individual-carbon-footprint-calculation?resource=download&select=Carbon+Emission.csv**](https://www.kaggle.com/datasets/dumanmesut/individual-carbon-footprint-calculation?resource=download&select=Carbon+Emission.csv)