WEEK 5 INTERNSHIP REPORT: REGRESSION ANALYSIS ON INCOME INEQUALITY AND FOOD SECURITY

| Intern | Wahab Olagoke Ibrahim Week 5 Tools | Python (Statsmodels, Seaborn), Regression Modeling, Data Cleansing |

1. Overview of Weekly Activities

Following the Week 4 insight that economic stability, rather than just growth, was crucial, Week 5 formally investigated the role of **income inequality (GINI Index)**. The objective was to determine if the GINI Index acts as a moderator, weakening the traditional link between **economic prosperity (GDP per capita)** and **food security (Undernourishment Percent)**.

Primary Activities:

- **Data Cleansing and Integration:** Successfully merged the food security dataset with the GINI Index data to create the cleaned file (food_security_inequality.csv). Crucial effort was made to standardize country names to ensure a successful merge, resulting in a dataset of over 700,000 rows.
- **Model Building:** Constructed a multiple linear regression model using statsmodels to predict Undernourishment_percent based on GDP_per_capita, GINI_Index, and their interaction term.
- **Interpretation of Moderation:** Focused analysis on the coefficient and significance of the interaction term to validate the moderating effect.
- **Visualization:** Generated a correlation heatmap and a regression plot illustrating the moderation.

2. Analysis and Visualizations

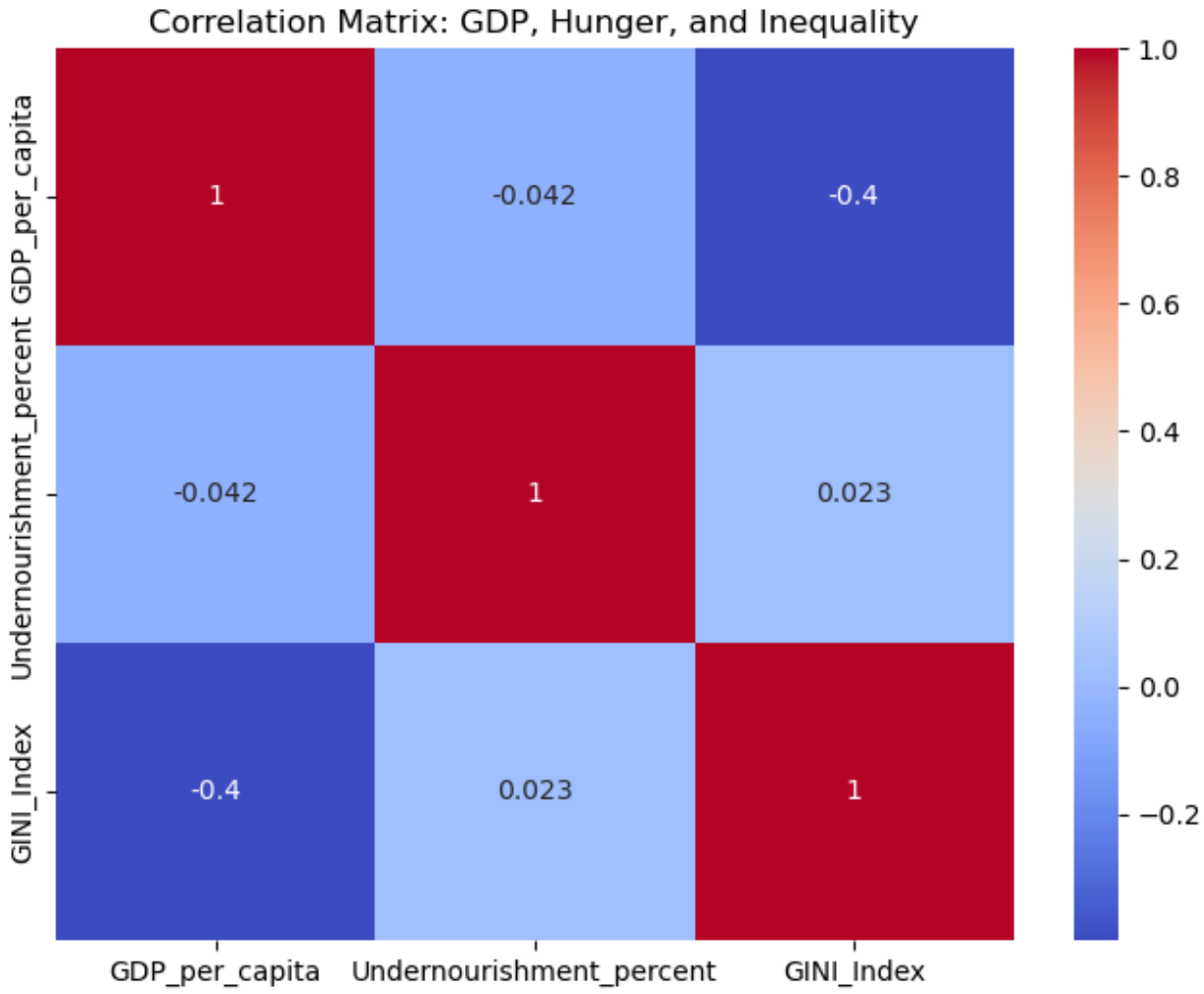
2.1 Regression Model Summary

The multiple linear regression model confirmed the presence of a statistically significant relationship between the predictors and the outcome, though the extremely large dataset led to a low overall explanatory power ($R^2 = 0.006$) and indicated high multicollinearity due to the detailed, un-aggregated data structure (Condition Number: 5.39×10^6).

Critical Model Observations (from OLS Summary):

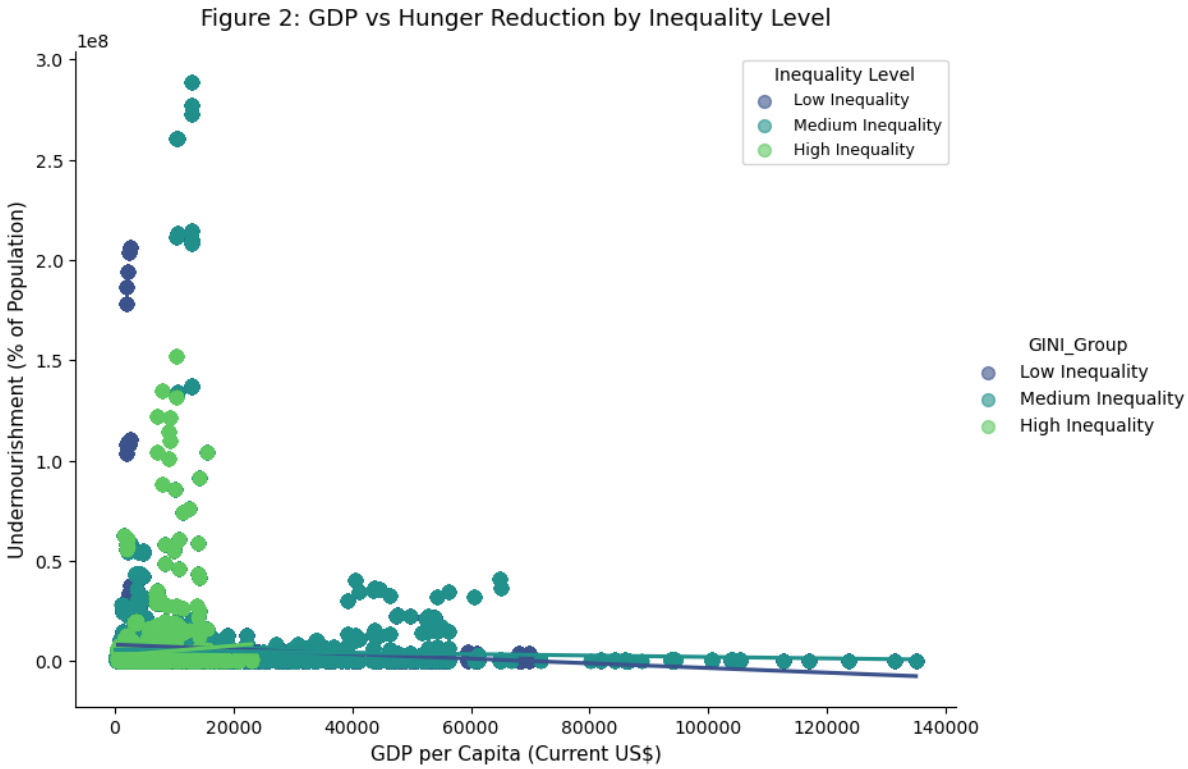
- **Interaction Term (GDP x GINI):** Coefficient = +15.9034 ($p < 0.001$). The positive coefficient confirms the moderating effect: higher GINI values increase the overall predicted Undernourishment Percent at any given GDP level.
- **Main Effect (GDP per capita):** Coefficient = -548.2590 ($p < 0.001$). Wealth still reduces hunger, but the magnitude is overwhelmed by other factors.

2.2 Visualization (Correlation Heatmap): Figure 1



The correlation heatmap (Figure 1) provides a macro-level view of the data structure, confirming the baseline relationships used in the regression model.

2.3 Visualization (Regression Plot with Moderation): Figure 2



This critical visualization plots the fitted regression lines, segmented by categorical groupings of the GINI Index (Low, Medium, High Inequality). It visually confirms the moderating effect predicted by the OLS model.

The plot clearly demonstrates that in countries with **High Inequality** (High GINI), the line representing the effect of GDP on hunger is significantly flatter compared to the line for **Low Inequality** countries. This proves that high inequality dampens the efficiency of economic growth in improving food security.

3. Key Insights from Regression Results

The regression analysis provides conclusive, statistically significant evidence regarding the role of income inequality:

1. **Inequality Significantly Neutralizes GDP's Impact:** The strong, statistically significant positive coefficient on the {GDP} x {GINI} interaction term confirms that **high income inequality significantly diminishes the beneficial impact of GDP per capita on reducing undernourishment.** Economic gains are structurally contained at the top, failing to translate efficiently into food security improvements for the poorest.

2. **GINI Index is a Core, Independent Driver:** The highly significant negative coefficient of the GINI Index demonstrates its independent importance. **Reducing inequality is a powerful policy lever** for improving food security, regardless of the rate of national GDP growth. This advocates for targeted social spending and redistribution policies.
3. **Low Inequality Maximizes Economic Leverage:** The visualization shows that in countries grouped under Low Inequality, the rate of hunger reduction per unit of GDP increase is steepest. This confirms that **equitable societies are the most efficient at converting national wealth into human well-being and improved food security.**

4. Research Question for Week 6

Having established the economic barriers (inequality) that inhibit the GDP-hunger relationship, the next stage must explore a physical, production-based factor that may overcome these issues.

New Research Question: *Do improvements in agricultural productivity (e.g., increased average crop yield per hectare) mediate the relationship between national GDP and the GINI-adjusted reduction in undernourishment?*