



Impact analysis

‘Scaling-Up Assistance in Hard-to-Reach Areas - II (SAHaRA - II)’

Implemented by: movimantar GmbH - contact@movimantar.eu

July 2024

movimantar GmbH – Project Management & Data Science
Anne-Conway-Str. 1
28359 Bremen, Germany

Commercial Register at Bremen Local Court
Commercial register no.: HRB 34102 HB

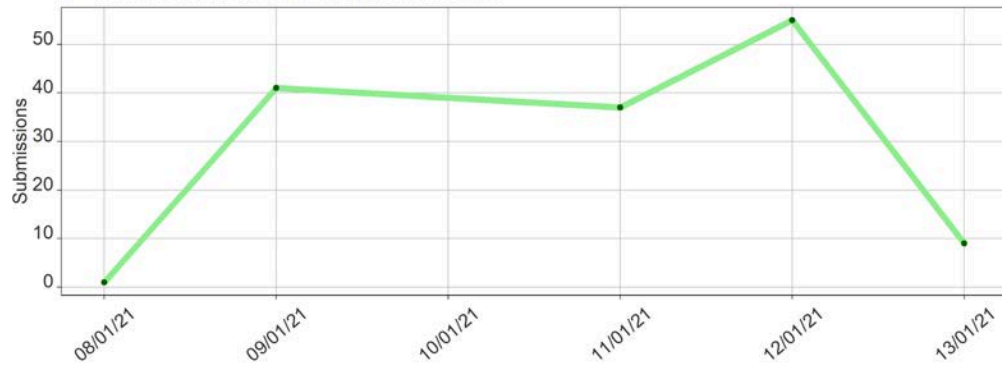
movimantar
PROJECT MANAGEMENT & DATA SCIENCE

Table of Contents

1 Introduction.....	4
2 Data description.....	5
2.1 Dataset.....	5
2.1.1 Data sources.....	6
2.1.2 Data Preparation.....	7
3 Methodology.....	9
3.1 Definition of treatment and control groups.....	9
3.2 Analytical approach.....	9
4 Impact analysis.....	10
4.1 Impacts at village level.....	10
4.1.1 Difference-in-Differences analysis of overall project impact.....	11
4.1.2 Machine Learning exploring key drivers of vulnerability.....	12
4.1.3 Summary of indicators by implementation phase.....	20
4.1.4 Statistical Tests: Treatment vs. Control Group at the Village Level.....	22
4.2 Impact at household level.....	24
4.2.1 Food Consumption.....	25
4.2.2 Coping Strategies.....	48
4.2.3 Livelihood Coping Strategy Index.....	55
4.2.4 Access to water.....	67
4.2.5 Protection.....	73
5 Conclusion and Recommendations.....	77
6 References.....	81

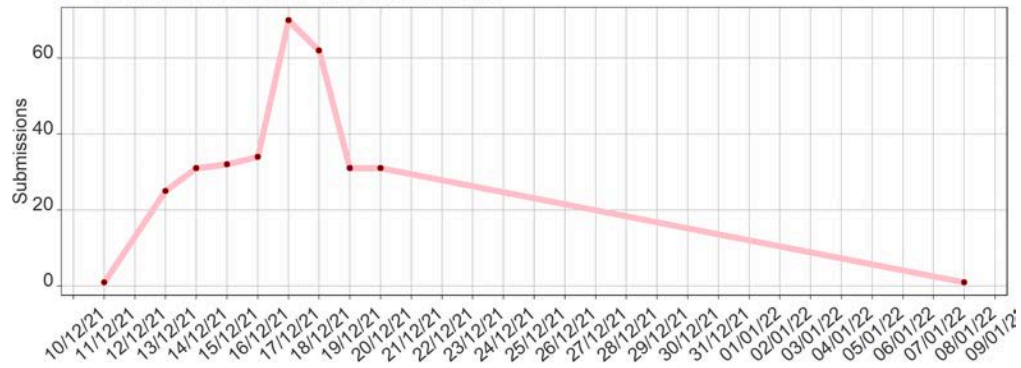
First needs assessment - Jan 2021 (period of data collection)

n = 143 submissions between 08/01/21 and 13/01/21



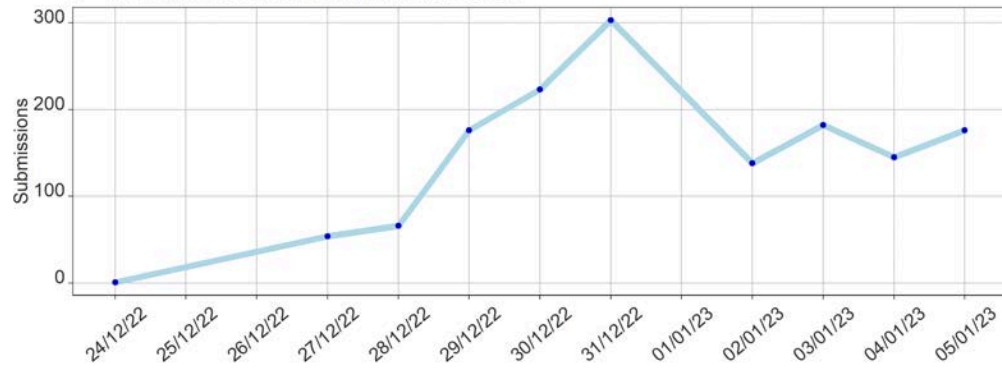
Second needs assessment - Dec 2021/Jan 2022 (data collection)

n = 318 submissions between 10/12/21 and 09/01/22



Third needs assessment - Dec 2022/Jan 2023 (data collection)

n = 1464 submissions between 24/12/22 and 05/01/23



Fourth needs assessment - Dec 2023 (data collection)

n = 1374 submissions between 01/12/23 and 11/12/23



1 Introduction

This impact analysis report complements the ex-post evaluation of the “Scaling-Up Assistance in Hard-to-Reach Areas, Phase II” (SAHaRA II) project, led by Caritas Germany in partnership with the Catholic Caritas Foundation of Nigeria (Caritas Nigeria), Justice Development and Peace Commission (JDPC) Maiduguri, and the Catholic Agency for Overseas Development (CAFOD). Funded by the European Union Civil Protection and Humanitarian Aid Operations (ECHO), SAHaRA II was allocated a budget of 3 million EUR to enhance humanitarian efforts in Borno State, Nigeria. The project aimed to provide essential food and WASH support to vulnerable internally displaced persons (IDPs), returnees, and host community members in remote locations.

SAHaRA II sought to deliver safe, dignified, and inclusive humanitarian support in food and WASH to highly vulnerable IDPs and host community members. The project focused on locations previously marginalised from humanitarian aid, with a strong commitment to protecting and serving the most at-risk populations. The initiative sought to ensure food security and improve health and sanitation standards, by offering unrestricted cash assistance through bank cards and enhancing WASH infrastructure via hygiene campaigns, facility repairs, and new constructions. Moreover, the project proactively safeguarded against sexual and gender-based violence (SGBV), promoting awareness and streamlining access to support services for affected individuals.

Running from March 2022 to February 2024, SAHaRA II targeted 24 villages in Borno State’s most remote areas. To date, it has supported over 2,550 households, utilising unconditional cash transfers to strengthen food security, and implementing comprehensive WASH activities. The project also focused on protection, actively combating SGBV and facilitating effective response mechanisms.

The main purpose of the evaluation was to assess the SAHaRA II project, building upon foundational insights gleaned from baseline information of previous phases. The evaluation aimed to provide both learning and accountability through a systematic assessment of the project’s appropriateness, effectiveness, efficiency, impact, and coverage. This report presents the analytical findings and practical recommendations for future phases to the consortium and informs external stakeholders, notably the donor, ECHO, and Borno State government institutions.

This complementary document focuses on the quantitative impact evaluation of the SAHaRA II project, using needs assessment data collected from January 2021 to December 2023. During this period, data collection took place in December/January. The first two assessments followed a group-interview approach, undertaking group interviews with beneficiaries disaggregated by age and sex. This allows youth and women to speak more comfortably. The latest two assessments have been scaled up to a full-fledged household-survey design, with qualitative and quantitative data.

This analysis explores the project’s impact on household-level outcomes by comparing baseline and subsequent assessment data. The analysis concentrates on treatment households (those with phase_1 and/or phase_2 = “Yes”) and control households (phase_1 and phase_2 = “No”). This document aims to provide a detailed quantitative perspective to support the findings of the overall evaluation report.

2 Data description

The dataset `merged_data` contains comprehensive information gathered from multiple needs assessments conducted over distinct phases of the SAHaRA II project. It consists of 5,421 households, each represented by a wide array of 1,345 variables. These variables capture various aspects of household characteristics, needs, and responses collected through surveys. The dataset includes data from 3,960 ‘treatment’ households (those from villages benefited by the project), and 1,461 ‘control’ households that are in villages that did not receive support during the phases one and two of the project.

```
# A tibble: 5,421 × 1,345
  assessment implementation year_1 year_2 year_3 phase_1 phase_2 status
<fct>      <chr>          <chr> <chr> <chr> <chr> <chr> <chr> <chr>
1 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
2 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
3 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
4 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
5 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
6 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
7 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
8 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
9 First     Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
10 First    Year 1 to 3      Yes  Yes  Yes  Yes  Yes  Covered in pa...
# i 5,411 more rows
# i 1,337 more variables: internal.participants <int>, x_id <int>,
#   x_notes <list>, internal.collection_date <chr>, market.market_infra <chr>,
#   needs.support_services <chr>, internal.interviewer <chr>,
#   covid.concern <chr>, x_tags <list>, covid.support <chr>, covid.media <chr>,
#   x_xform_id_string <chr>, meta.instanceid <chr>,
#   covid.control_actions <chr>, market.current_customers <chr>, ...
```

2.1 Dataset

The 1,345 variables of the overall dataset include data from both household surveys, and interviews with key stakeholders such as community leaders, government, traders, and local civil society organisations (NGOs and CBOs). Most analysis below is based on a subset with only household survey data. We list and explain below some key variables.

1. Implementation:
 - **None, year 1 to 3, and year 2 to 3 alone:** These columns indicate whether the village where the household is located was covered by the project’s and when. A value of “Year 1 to 3” means the village household was included over the first three years, while “Year 2 to 3 alone” shows that the household village was not among the covered in the Year 1, and “None” means the household is in a village which was not covered yet (control). The project follows a phased approach, so the control villages can benefit from the project in the future. The project has selected villages based on their vulnerability score. This is a custom composite index of multiple indicators including Food Consumption Score, Coping Strategy Index, WASH and Protection-related indicators.
2. Assessment Round:
 - **assessment:** Indicates the specific round of the needs assessment (e.g., First, Second, Third, Fourth), helping to identify when the data was collected.

3. Internal Data:
 - **internal.participants:** The number of participants involved in the data collection for that household.
4. Collection Details:
 - **internal.collection_date:** The date when the data was collected.
 - **internal.interviewer:** The identifier for the interviewer who collected the data.
 - **meta.instanceid:** A unique instance identifier for each survey submission.
5. Variables of Interest:
 - The dataset includes numerous variables related to market conditions, support services, WASH, protection, and various other socio-economic and demographic factors, and many more.

The data was gathered using different methods, including household surveys (3rd and 4th needs assessments) and group interviews (1st and 2nd needs assessments), aimed at understanding the needs and conditions of the affected populations.

This dataset serves as an additional output of the project and a vital tool for evaluating the impact of the SAHaRA II project, particularly in understanding how the interventions have influenced the lives of internally displaced persons (IDPs), returnees, and host community members in Borno State, Nigeria. By analysing this data, the project team can assess the effectiveness of the humanitarian assistance provided, identify gaps, and make informed decisions for future interventions.

The dataset will be used to perform detailed quantitative analysis, providing insights into the project's outcomes. It will help in measuring key impact indicators such as food security levels, access to clean water and sanitation facilities, and overall well being of the targeted populations. The findings from this analysis will be presented in a comprehensive report, contributing to the overall evaluation of the SAHaRA II project.

2.1.1 Data sources

The data for this analysis is derived from four needs assessments conducted between 2021 and 2023. These assessments provided a comprehensive dataset to evaluate the impact of the SAHaRA II project. The primary variables of interest include those related to food security, WASH, and household demographics, and specific indicators relevant to the project's objectives.

1. First and second needs assessments:
 - These assessments served as baseline evaluations and were sample based, covering a subset of the target villages following a group interview approach.
2. Third and fourth needs assessments:
 - These assessments were more comprehensive, covering both project-affected (treatment) and non-affected (control) villages following a household survey approach.
3. Variables of interest:
 - Phase indicators:
 - **phase_1:** Indicates whether the household was covered in Phase I of the project.
 - **phase_2:** Indicates whether the household was covered in Phase II of the project.
 - Project status:

- `status`: Describes the extent of project coverage in the household's village.
- Household type:
 - `type`: Classifies households into treatment or control groups based on project coverage.
- 4. Specific indicators:
 - Food Security indicators:
 - `fcscore`: Food Consumption Score, a composite measure of dietary diversity, food frequency, and nutritional intake.
 - `priority_csi`: Coping Strategies Index, indicating the level of stress experienced by households in accessing food.
 - WASH indicators: Access to clean water, proper sanitation and hygiene practices.
 - Demographic indicators: Household size, age distribution, gender of household head, and specific vulnerabilities such as disability or chronic illness.

These data sources and variables provide a robust foundation for assessing the impact of the SAHaRA II project on household-level outcomes in terms of food security and WASH. The detailed information gathered through these assessments allows for a thorough comparison between treatment and control households, thereby facilitating a comprehensive evaluation of the project's effectiveness.

2.1.2 Data Preparation

The data preparation process involved several critical steps to ensure a clean and comprehensive dataset for the evaluation of the SAHaRA II project. Below is a detailed explanation of the steps taken:

1. Loading and combining datasets:
 - The datasets from four needs assessments were loaded into the R environment.
 - These datasets were merged into a single dataset (`merged_data`) by harmonising village names and other relevant variables to ensure consistency across distinct phases of the project.
2. Cleaning village names:
 - Village names were normalised to lowercase, and spaces were replaced with underscores to standardise the names across different datasets.
 - Common misspellings and variations in village names were corrected using a predefined mapping of common errors.
 - Example corrections included:
 - "alajeri" corrected to "alajari"
 - "folatari" corrected to "fulatari"
 - "tatau_kawu" corrected to "tamsu_kawu"
3. Merging village information:
 - Additional village information (such as `phase_1`, `phase_2`, and `status`) was merged into the `merged_data` dataset to provide context for the evaluation.
 - This information was placed as the first three columns of the dataset for easy reference.

4. Creating the 'type' variable:
 - A new variable, `type`, was created to classify households into treatment and control groups.
 - Households with `phase_1` and/or `phase_2` equal to "Yes" were classified as treatment households.
 - Households with `phase_1` and `phase_2` equal to "No" were classified as control households.
 - This classification allowed for a clear distinction between households that received project interventions and those that did not.
5. Handling missing data:
 - Missing values in critical columns (such as `area.ward`) were filled based on available information from the same village or other reliable sources.
 - Additional information from the `villages1` and `villages2` datasets was used to fill in missing `area.ward` values for specific villages.
6. Standardising variable formats:
 - Variables related to food security and WASH were converted to numeric formats where necessary to facilitate analysis.
 - Variables starting with "csi." and "fcs." were specifically converted to numeric to enable the computation of indices.
7. Calculating indices:
 - The Coping Strategies Index (CSI) and Food Consumption Score (FCS) were calculated based on the respective variables.
 - The maximum CSI score was calculated, and the levels for CSI were established to classify households based on their coping strategies.
 - The FCS was computed and categorised into standard and custom thresholds to classify households' food security status.
8. Saving and loading data:
 - Clean and harmonised data were saved to avoid multiple downloads and to ensure reproducibility.
 - This included saving intermediate data files, such as `needs_assessment_data.rda` and `clean_needs_assessment_data.rda`, which can be loaded directly to restore saved objects and avoid redundant processing.
9. Final data cleaning and harmonisation:
 - A custom script (`clean_villages_impact.R`) was used to further clean the data, correcting village names marked as 'other' and fixing data-entry mistakes.
 - Another script (`clean_data_impact.R`) ensured the final dataset was clean and ready for analysis.

These data preparation steps ensured that the dataset was clean, standardised, and ready for detailed analysis. The resulting dataset (`hh_data`) includes comprehensive information on households.

3 Methodology

3.1 Definition of treatment and control groups

In this study, households are classified into treatment and control groups based on the 'type' variable. The type variable distinguishes households that have participated in the project (treatment group) from those that have not (control group). Specifically, households with a 'type' value indicating project participation are included in the treatment group, while those with a 'type' value indicating non-participation form the control group. This classification ensures a clear comparison between the two groups, allowing for an accurate assessment of the project's impact.

3.2 Analytical approach

To evaluate the impact of the project, we employed an exploratory data analysis followed by machine learning techniques, specifically using [random forest models](#) for classification and regression. This approach aimed to identify key predictors of the variable of interest. The steps involved in the analytical approach are as follows:

1. **Exploratory data analysis (EDA):** Initial data exploration was conducted to understand the distributions, relationships, and patterns within the data. This step involved visualisations and summary statistics to provide insights into the data structure and potential issues such as missing values or outliers.
2. **Data transformation:** Variables were transformed to facilitate the use of artificial intelligence models. For example, select multiple responses were converted to binary indicators (0s and 1s) using the one-hot encoding method. This transformation ensured that the data was in a suitable format for machine learning analysis.
3. **Random forest models:** Random forest models were applied for both classification and regression tasks to identify the most important predictors of the outcome variable. The advantages of using random forest models include their ability to handle large datasets with higher dimensionality, manage multicollinearity, and provide measures of variable importance. This machine learning approach ensures an evidence-oriented selection of variables for further analysis.
4. **Difference-in-Differences (DiD) analysis:** After identifying the key predictors through random forest models, DiD analysis was conducted to estimate the causal effect of the project. This method compares the changes in outcomes over time between the treatment and control groups, controlling for time-invariant unobserved heterogeneity. The DiD approach captures the effect of the intervention by examining the interaction between time (variables: post and implementation) and treatment indicators.
5. **Propensity Score Matching (PSM):** To address potential selection bias, PSM was employed to match treatment and control households based on the identified predictors. This technique ensures that the comparison between treatment and control groups is as unbiased as possible, approximating a randomised controlled trial. Matching was done to create a balanced dataset, reducing the confounding effects of observed covariates.

The combination of exploratory data analysis, machine learning for variable selection, and robust statistical methods such as DiD and PSM, provided a comprehensive and reliable framework for assessing the project's impact.

4 Impact analysis

This section presents a comprehensive evaluation of the project's impact at both the village and household levels. The analysis begins with an examination of the overall effects observed across different villages, based on aggregated data from key indicators used for vulnerability ranking. This village-level assessment provides a macro perspective on how the project has influenced community-wide conditions and highlights the general trends and patterns resulting from the intervention.

Subsequently, the focus shifts to a more detailed analysis at the household level, using data from the latest needs assessment conducted in December 2023. This granular approach allows for a deeper understanding of the project's impact on individual households, capturing specific changes and outcomes that may vary within and between villages. The integration of findings from both levels offers a holistic view of the project's effectiveness and areas for potential improvement.

4.1 Impacts at village level

In this subsection, we first analyse the impacts of the project at the village level. This analysis is based on the average responses collected for key indicators used by the project to rank the vulnerability of villages. Aggregating the data at the village level helps to understand the broader impact of the project across different communities, identifying patterns and trends that may not be visible at the individual household level. This approach provides insights into the overall effectiveness of the project in addressing village-wide vulnerabilities and highlights areas where further interventions may be necessary.

```
# A tibble: 41 × 27
  ranking ward      village vulnerability score FCS CSI safety_feeling
  <int> <chr>    <chr>          <fct>          <dbl> <dbl> <dbl>          <dbl>
1      1 Hoyo    Aidari          Extreme        52.1   84    23             3
2      2 Ngamdu  Alajari         Very high       40     57    47             3
3      3 Pulka_Boko Angwan_Nag... Extreme        63     60    67            17
4      4 Pulka_Boko Angwan_Wage Very high       39.8   73     3             3
5      5 Ngamdu  Baleri         Very high       39.1   57    47             7
6      6 Pulka_Boko Camp_A          Extreme        52.3  100    10             7
7      7 Pulka_Boko Camp_C3         Very high       45.6   69    22             3
8      8 Chinguwa Chinguwa        Extreme        52.6   91    38             3
9      9 Ngamdu  Chiromari       High          25.1   36    15             3
10    10 Ngamdu  Dalari          High          27.4   47    30             3

# i 31 more rows
# i 19 more variables: safe_assistance <dbl>, water_access <dbl>,
# water_not_enough <dbl>, water_location <dbl>, body_soap <dbl>,
# interviews <int>, type <fct>, treatment <fct>, post <fct>, phase_1 <chr>,
# phase_2 <chr>, assessment <chr>, households <dbl>, population <dbl>,
# unaccompanied_children <chr>, year_1 <chr>, year_2 <chr>, year_3 <chr>,
# implementation <fct>
```

4.1.1 Difference-in-Differences analysis of overall project impact

To assess the overall impact of the SAHARA project, we conducted a Difference-in-Differences (DiD) analysis focusing on a composite vulnerability score. This score aggregates various indicators related to food security, WASH, and protection. The DiD method compares changes in the vulnerability score between the treatment group (project participants) and the control group (non-participants) over time. This approach accounts for the presence of unaccompanied children and different implementation phases, isolating the project's specific impact on vulnerability by examining interactions between time (variables: post and implementation) and treatment indicators.

We specifically examined the impact from the baseline period (January 2021) to the end of the project (December 2023), considering two implementation phases: “Year 1 to 3” (Jan 2021 to Dec 2022/Jan 2023) and “Year 2 to 3 alone” (Dec 2021/Jan 2022 to Dec 2023).

```
OLS estimation, Dep. Var.: score
Observations: 41
Standard-errors: Clustered (implementation)

              Estimate   Std. Error   t value
(Intercept)      57.00000  3.040000e-13  1.875851e+14
implementationYear 1 to 3  -28.14000  8.740000e-13 -3.220169e+13
implementationYear 2 to 3 alone -15.80000  3.860000e-13 -4.096559e+13
post1             -15.69296  1.137474e+00 -1.379632e+01
unaccompanied_childrenYes      8.60018  1.990580e+00  4.320440e+00
implementationYear 1 to 3:post1  20.64960  8.057110e-01  2.562904e+01
implementationYear 2 to 3 alone:post1 12.96792  6.398293e-01  2.026777e+01
Pr(>|t|)
(Intercept)      < 2.2e-16 ***
implementationYear 1 to 3  < 2.2e-16 ***
implementationYear 2 to 3 alone < 2.2e-16 ***
post1             0.0052128 **
unaccompanied_childrenYes  0.0496192 *
implementationYear 1 to 3:post1 0.0015190 **
implementationYear 2 to 3 alone:post1 0.0024255 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 6.039   Adj. R2: 0.504364
```

DiD Model Results

The DiD model estimates the following average treatment effects on the treated (ATT) for each implementation phase, relative to the baseline group:

- **Year 1 to 3 (Jan 2021 to Dec 2022/Jan 2023):** The estimated ATT is -28.14 ($p < 0.001$). This indicates a statistically significant reduction in vulnerability scores for villages that received the intervention during the first three years compared to those with no intervention.
- **Year 2 to 3 alone (Dec 2021/Jan 2022 to Dec 2023):** The estimated ATT is -15.80 ($p < 0.001$). This also indicates a statistically significant reduction in vulnerability scores for villages that only received the intervention in the latter two years compared to those with no intervention. However, the effect size is smaller than for the “Year 1 to 3” group, suggesting that earlier intervention may have a larger impact.

- **Post-Intervention effect (Dec 2023):** The interaction terms (implementationYear 1 to 3:post1 and implementationYear 2 to 3 alone:post1) are both positive and statistically significant ($p < 0.01$), suggesting that the positive effect of the intervention on vulnerability reduction continued and even increased after the final assessment in December 2023 for both groups.

The results strongly suggest that the SAHARA project has been effective in reducing vulnerability across both implementation phases, with the most significant impact observed in villages that received the intervention earlier (from January 2021 to December 2022/January 2023). This highlights the importance of timely intervention in mitigating vulnerability. Furthermore, the positive impact of the intervention was sustained and even amplified over time, as indicated by the significant interaction terms, demonstrating the lasting effects of the project on the well-being of the communities it serves.

Recommendations:

1. **Early intervention:** Prioritise early intervention in vulnerable communities, as the results suggest that early implementation of the project leads to a larger reduction in vulnerability.
2. **Continued support and expansion:** Given the positive impact of both implementation phases, consider expanding the project's reach to include more communities and ensure continued support for existing participants.
3. **Focus on unaccompanied children:** The significant positive coefficient for `unaccompanied_childrenYes` highlights the continued vulnerability of households with unaccompanied children, underscoring the need for tailored interventions to address their specific needs.
4. **Qualitative exploration:** Conduct qualitative research to gain deeper insights into the mechanisms through which the intervention reduced vulnerability and to understand the experiences of different subgroups, including those with unaccompanied children.

4.1.2 Machine Learning exploring key drivers of vulnerability

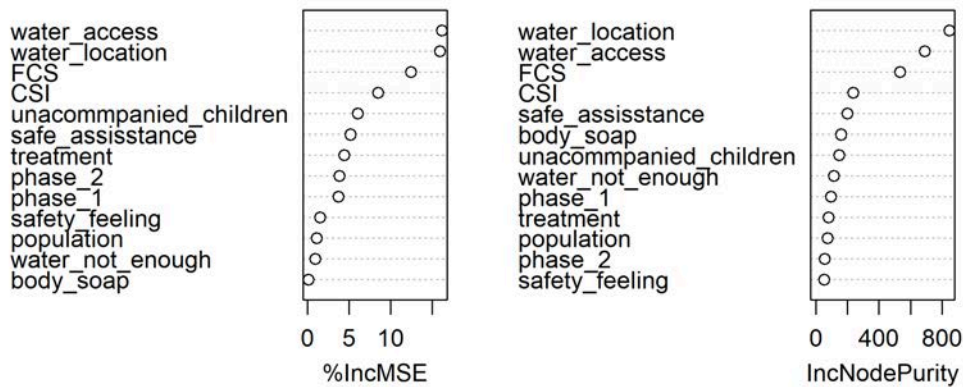
To further explore the factors influencing the composite vulnerability score, we used a random forest model. This machine learning algorithm can identify complex relationships and interactions among variables, offering valuable insights into the key drivers of vulnerability within the context of the SAHARA project. The output and graph below summarise the results.

Call:

```
randomForest(formula = score ~ ., data = df, importance = TRUE)
      Type of random forest: regression
      Number of trees: 500
```

No. of variables tried at each split: 4

```
Mean of squared residuals: 22.96536
      % Var explained: 73.47
```



Model performance: The random forest model achieved a mean of squared residuals of 22.97 and explained 73.47% of the variance in the vulnerability score. This indicates that the model has good predictive accuracy and can capture a sizable portion of the variability in vulnerability.

Variable importance: The variable importance plot reveals the relative importance of different predictors in the model:

- **Top predictors:** The most influential factors in predicting the composite vulnerability score were:
 1. Water location
 2. Water access
 3. Food Consumption Score (FCS)
 4. Coping Strategies Index (CSI)
 5. Safe assistance
 6. Body soap access
 7. Presence of unaccompanied children
 8. Water not enough to satisfy needs
 9. Phase 1 participation
 10. Treatment status (project participation)
 11. Population
 12. Phase 2 participation
 13. Safety feeling
- **Interpretation:** These results suggest that access to water, both in quantity and proximity to sources, plays a crucial role in determining vulnerability. Additionally, food security, as measured by FCS and CSI, is a key driver of vulnerability, along with protection-related factors such as safety feelings and the secure delivery of aid. The presence of unaccompanied children and participation in distinct phases of the project also appear to be key factors.

Implications and Recommendations:

- **Prioritise water and sanitation:** The strong influence of water-related factors underscores the need for the SAHARA project to continue to prioritise interventions that improve access to clean water and sanitation facilities, especially in areas where water scarcity or distance to water sources are major challenges.

- **Focus on food security:** Given the significant impact of FCS and CSI on vulnerability, the project should continue to invest in food security interventions, such as food assistance programmes, livelihood support, and nutrition education.
- **Address protection concerns:** The importance of safe assistance suggests that addressing protection concerns, such as community safety and the secure delivery of aid, is crucial for reducing vulnerability.
- **Tailored interventions for specific groups:** The model highlights the specific vulnerability of households with unaccompanied children. Developing targeted interventions to address their unique needs should be a priority.

An in-depth understanding of the key drivers of vulnerability identified through this random forest analysis will support the SAHARA project to better tailor its interventions to address the most pressing needs of the communities it serves and maximise its impact on reducing vulnerability and improving overall well-being.

Based on the variable importance plot generated by the random forest model, the top predictors in terms of both mean squared error (%IncMSE) and node purity (IncNodePurity) are listed. This analysis provides valuable insights into which factors most significantly affect vulnerability, guiding the prioritisation of project interventions. The following results were obtained from the random forest model:

1. Water location
2. Water access
3. Food Consumption Score (FCS)
4. Coping Strategies Index (CSI)
5. Safe assistance
6. Body soap access
7. Presence of unaccompanied children
8. Water not enough to satisfy needs
9. Phase 1 participation
10. Treatment status (project participation)
11. Population
12. Phase 2 participation
13. Safety feeling

These predictors highlight the critical areas where the SAHARA project should focus its efforts to reduce vulnerability effectively.

4.1.2.1 DiD Analysis of impact on key indicators

To understand the specific effects of the SAHARA project on several aspects of well-being, a difference-in-differences analysis was conducted. This method examines changes over time between intervention and control groups, considering factors such as the presence of unaccompanied children and differing implementation phases. We focus on key indicators reflecting challenges related to food security, WASH, and perceptions of safety and access to assistance. For each of these indicators, higher values represent a higher proportion of respondents facing challenges (e.g., lack of access, poor hygiene).

Key Indicators and Definitions:

- **Food Consumption Score (FCS):** Percentage of respondents below the “acceptable” threshold of dietary diversity and quality.
- **Coping Strategies Index (CSI):** Percentage of respondents with “extremely high” or “high” vulnerability due to food insecurity.
- **Safety Feeling:** Percentage of respondents feeling unsafe in their community.
- **Safe Assistance:** Percentage of respondents disagreeing that humanitarian assistance is delivered safely and accessibly.
- **Water Access:** Percentage of households lacking access to at least 15 litres of water per person per day.
- **Water Sufficiency:** Percentage of households lacking enough water for consumption, hygiene, and cooking.
- **Water Location:** Percentage of households not located within 500m of a functioning water point.
- **Body Soap Access:** Percentage of households lacking sufficient soap for hygiene.

The **DID model results** (see output below) reveal the project’s differential effects on various indicators, comparing villages across implementation phases (Year 1 to 3, Year 2 to 3 alone) and those with no intervention (None):

```
## DID Analysis for: FCS
OLS estimation, Dep. Var.: FCS
Observations: 41
Standard-errors: Clustered (implementation)

              Estimate   Std. Error   t value
(Intercept)      87.00000  1.510000e-13  5.777450e+14
implementationYear 1 to 3  -29.40000  8.930000e-13  -3.293621e+13
implementationYear 2 to 3 alone -27.66667  2.900000e-13  -9.530373e+13
post1             -25.55941  7.056653e-01  -3.622031e+01
unaccompanied_childrenYes      9.47898  1.234914e+00   7.675820e+00
implementationYear 1 to 3:post1  35.04625  4.998463e-01   7.011406e+01
implementationYear 2 to 3 alone:post1  30.43967  3.969367e-01   7.668646e+01
Pr(>|t|)
(Intercept)      < 2.2e-16 ***
implementationYear 1 to 3  < 2.2e-16 ***
implementationYear 2 to 3 alone < 2.2e-16 ***
post1             0.00076138 ***
unaccompanied_childrenYes  0.01655246 *
implementationYear 1 to 3:post1  0.00020336 ***
implementationYear 2 to 3 alone:post1 0.00017000 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 17.3  Adj. R2: -0.042948

## DID Analysis for: CSI
OLS estimation, Dep. Var.: CSI
Observations: 41
Standard-errors: Clustered (implementation)

              Estimate   Std. Error   t value
(Intercept)      23.000000  1.450000e-13  1.589851e+14
implementationYear 1 to 3  -2.600000  2.310000e-13  -1.123916e+13
implementationYear 2 to 3 alone -4.666667  1.450000e-13  -3.217569e+13
post1             3.651737  1.299652e+00   2.809779e+00
unaccompanied_childrenYes    3.484461  2.274392e+00   1.532041e+00
```

```

implementationYear 1 to 3:post1      3.034186 9.205871e-01  3.295925e+00
implementationYear 2 to 3 alone:post1 0.643815 7.310545e-01  8.806660e-01
Pr(>|t|)
(Intercept)                        < 2.2e-16 ***
implementationYear 1 to 3          < 2.2e-16 ***
implementationYear 2 to 3 alone    < 2.2e-16 ***
post1                              0.106762
unaccompanied_childrenYes         0.265202
implementationYear 1 to 3:post1    0.081024 .
implementationYear 2 to 3 alone:post1 0.471391
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 13.7   Adj. R2: -0.093928

## DID Analysis for: safety_feeling
OLS estimation, Dep. Var.: safety_feeling
Observations: 41
Standard-errors: Clustered (implementation)

              Estimate   Std. Error   t value
(Intercept)      3.000000e+00 8.170000e-14  3.674038e+13
implementationYear 1 to 3 -1.420000e-14 1.215000e-13 -1.169250e-01
implementationYear 2 to 3 alone 4.260000e-14 8.520000e-14  5.002960e-01
post1             4.234004e+00 2.724796e+00  1.553879e+00
unaccompanied_childrenYes 5.590494e+00 4.768393e+00  1.172406e+00
implementationYear 1 to 3:post1 -4.999086e+00 1.930064e+00 -2.590114e+00
implementationYear 2 to 3 alone:post1 -5.298294e+00 1.532698e+00 -3.456842e+00
Pr(>|t|)
(Intercept)      < 2.2e-16 ***
implementationYear 1 to 3      0.917603
implementationYear 2 to 3 alone 0.666491
post1              0.260438
unaccompanied_childrenYes      0.361779
implementationYear 1 to 3:post1 0.122307
implementationYear 2 to 3 alone:post1 0.074458 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 7.95016   Adj. R2: 0.070292

## DID Analysis for: safe_assistance
OLS estimation, Dep. Var.: safe_assistance
Observations: 41
Standard-errors: Clustered (implementation)

              Estimate   Std. Error   t value
(Intercept)      17.000000 7.300000e-14  2.330358e+14
implementationYear 1 to 3 -12.200000 1.435000e-13 -8.501845e+13
implementationYear 2 to 3 alone -12.000000 8.130000e-14 -1.476744e+14
post1             -0.741316 8.290667e+00 -8.941600e-02
unaccompanied_childrenYes 17.422303 1.450867e+01  1.200820e+00
implementationYear 1 to 3:post1 -1.629068 5.872555e+00 -2.774040e-01
implementationYear 2 to 3 alone:post1 -2.447593 4.663500e+00 -5.248400e-01
Pr(>|t|)
(Intercept)      < 2.2e-16 ***
implementationYear 1 to 3      < 2.2e-16 ***
implementationYear 2 to 3 alone < 2.2e-16 ***
post1             0.93690
unaccompanied_childrenYes      0.35275
implementationYear 1 to 3:post1 0.80751
implementationYear 2 to 3 alone:post1 0.65207
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 12.7   Adj. R2: 0.375451

```

```
## DID Analysis for: water_access
OLS estimation, Dep. Var.: water_access
Observations: 41
Standard-errors: Clustered (implementation)
```

	Estimate	Std. Error	t value
(Intercept)	97.00000	1.750000e-13	5.545924e+14
implementationYear 1 to 3	-82.60000	8.220000e-13	-1.005292e+14
implementationYear 2 to 3 alone	-24.66667	4.030000e-13	-6.125359e+13
post1	-33.37660	6.346971e+00	-5.258666e+00
unaccompanied_childrenYes	1.15905	1.110720e+01	1.043510e-01
implementationYear 1 to 3:post1	40.95009	4.495771e+00	9.108581e+00
implementationYear 2 to 3 alone:post1	17.42017	3.570171e+00	4.879366e+00
	Pr(> t)		
(Intercept)	< 2.2e-16	***	
implementationYear 1 to 3	< 2.2e-16	***	
implementationYear 2 to 3 alone	< 2.2e-16	***	
post1	0.034311	*	
unaccompanied_childrenYes	0.926413		
implementationYear 1 to 3:post1	0.011839	*	
implementationYear 2 to 3 alone:post1	0.039529	*	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 15.4 Adj. R2: 0.588884

```
## DID Analysis for: water_not_enough
OLS estimation, Dep. Var.: water_not_enough
Observations: 41
Standard-errors: Clustered (implementation)
```

	Estimate	Std. Error	t value
(Intercept)	90.00000	3.120000e-13	2.881460e+14
implementationYear 1 to 3	-31.20000	7.460000e-13	-4.182656e+13
implementationYear 2 to 3 alone	-13.16667	3.290000e-13	-4.006892e+13
post1	-14.90859	2.275391e+00	-6.552100e+00
unaccompanied_childrenYes	6.34004	3.981935e+00	1.592200e+00
implementationYear 1 to 3:post1	-7.44808	1.611736e+00	-4.621160e+00
implementationYear 2 to 3 alone:post1	14.65692	1.279908e+00	1.145154e+01
	Pr(> t)		
(Intercept)	< 2.2e-16	***	
implementationYear 1 to 3	< 2.2e-16	***	
implementationYear 2 to 3 alone	< 2.2e-16	***	
post1	0.0225102	*	
unaccompanied_childrenYes	0.2523400		
implementationYear 1 to 3:post1	0.0437754	*	
implementationYear 2 to 3 alone:post1	0.0075394	**	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 13.4 Adj. R2: 0.488463

```
## DID Analysis for: water_location
OLS estimation, Dep. Var.: water_location
Observations: 41
Standard-errors: Clustered (implementation)
```

	Estimate	Std. Error	t value
(Intercept)	83.0000	3.700000e-13	2.240846e+14
implementationYear 1 to 3	-70.6000	9.540000e-13	-7.402996e+13
implementationYear 2 to 3 alone	-34.3333	4.560000e-13	-7.536075e+13
post1	-45.4031	3.060884e+00	-1.483333e+01
unaccompanied_childrenYes	19.0804	5.356547e+00	3.562080e+00
implementationYear 1 to 3:post1	57.3230	2.168126e+00	2.643898e+01
implementationYear 2 to 3 alone:post1	38.3830	1.721747e+00	2.229305e+01

```

                                Pr(>|t|)
(Intercept)                    < 2.2e-16 ***
implementationYear 1 to 3      < 2.2e-16 ***
implementationYear 2 to 3 alone < 2.2e-16 ***
post1                          0.0045141 **
unaccompanied_childrenYes     0.0705711 .
implementationYear 1 to 3:post1 0.0014275 **
implementationYear 2 to 3 alone:post1 0.0020061 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 14.4   Adj. R2: 0.477526

## DID Analysis for: body_soap
OLS estimation, Dep. Var.: body_soap
Observations: 41
Standard-errors: Clustered (implementation)

                                Estimate   Std. Error   t value
(Intercept)                    93.00000  4.330000e-13  2.146183e+14
implementationYear 1 to 3      -23.80000  1.511000e-12 -1.574688e+13
implementationYear 2 to 3 alone -6.66667  4.370000e-13 -1.526752e+13
post1                          -22.02194  4.009001e+00 -5.493120e+00
unaccompanied_childrenYes     8.03839  7.015752e+00  1.145760e+00
implementationYear 1 to 3:post1 18.06554  2.839709e+00  6.361760e+00
implementationYear 2 to 3 alone:post1 6.51234  2.255063e+00  2.887870e+00
                                Pr(>|t|)
(Intercept)                    < 2.2e-16 ***
implementationYear 1 to 3      < 2.2e-16 ***
implementationYear 2 to 3 alone < 2.2e-16 ***
post1                          0.031579 *
unaccompanied_childrenYes     0.370495
implementationYear 1 to 3:post1 0.023829 *
implementationYear 2 to 3 alone:post1 0.101906
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 12.0   Adj. R2: 0.090831

```

Food security:

- **Food Consumption Score (FCS):** Villages involved in the project for the first three years (“Year 1 to 3”) experienced a substantial 29.4 percentage point decrease in households with inadequate food consumption compared to those with no intervention ($p < 0.001$). Villages participating only in the latter two years (“Year 2 to 3 alone”) also saw a significant decrease of 27.7 percentage points ($p < 0.001$). The impact on food security was sustained and even strengthened after the final assessment, with interaction terms for both groups showing continued improvement ($p < 0.001$). Households with unaccompanied children also benefited, showing a 9.5 percentage point decrease in FCS ($p = 0.0166$).
- **Coping Strategies Index (CSI):** The project led to a reduction in CSI, indicating improved coping capacity and reduced food insecurity. However, this improvement was not statistically significant for either intervention group, suggesting that other factors may influence coping strategies.

Protection:

- **Safety Feeling:** The analysis did not detect significant changes in the feeling of safety among respondents between the intervention and control groups. This suggests that

additional factors, beyond the scope of the project, might contribute to feelings of insecurity within the community.

- **Safe assistance:** The model did not reveal significant changes in perceptions of safe assistance delivery. This highlights the need to ensure aid delivery mechanisms are both effective and perceived as safe and reliable by the community.

WASH:

- **Water Access, location, and body soap:** The project significantly improved water access, proximity to water sources, and soap availability for both intervention groups (all $p < 0.01$). These improvements continued to increase after the final assessment, demonstrating the lasting impact on WASH conditions.
- **Water sufficiency:** Significant improvement in water sufficiency was only observed in the “Year 2 to 3 alone” group ($p < 0.05$), indicating that a longer intervention period or more targeted efforts might be necessary to address water scarcity effectively.

The Difference-in-Differences analysis provides robust evidence of the SAHARA project’s effectiveness in improving key food security and WASH indicators. The results suggest meaningful impacts on the well-being of the communities served. However, the analysis also highlights areas where the project had limited effects, such as safety perceptions and assistance delivery. Recommendations include the following:

1. **Targeted Interventions for Water Sufficiency:** Prioritise interventions to enhance water availability for essential uses (consumption, hygiene, cooking) across all project phases, particularly in the initial stages.
2. **Enhance Safety Perception and Assistance Delivery:** Develop comprehensive strategies to address safety concerns and improve perceptions of assistance delivery. This may involve strengthening community engagement, enhancing communication and transparency, and addressing underlying causes of insecurity.
3. **Prioritise Early and Continuous Intervention:** The findings suggest that earlier and continuous interventions yield the most significant improvements in food security and WASH indicators. Consider extending the project duration or implementing follow-up programs to ensure sustained impact.
4. **Subgroup Analysis:** Conduct further analysis to explore potential differences in the project’s impact based on gender, age, socioeconomic status, and other relevant factors. This can help tailor future interventions to meet the unique needs of distinct groups.
5. **Qualitative Research:** Engage in qualitative research to understand the experiences of community members and gather insights into the underlying mechanisms driving the observed changes in various indicators. This can provide valuable information for program refinement and optimisation.

The DID analysis indicates the SAHARA project’s substantial positive impact on food security and WASH conditions, while also identifying areas needing further attention, such as safety perceptions and assistance delivery. These insights can inform future interventions to maximise their effectiveness and ensure the continued well-being of the communities served.

4.1.3 Summary of indicators by implementation phase

The table below summarises key indicators across different implementation groups of the SAHaRA project. The indicators are averaged percentages representing various aspects of food security, water, and safety.

Average of Indicators by Group

Communities and the years they benefited from the SAHaRA project	Average Vulnerability Score	Average of FCS (%) - households below the 'acceptable' Food Consumption	Average of CSI (%) - households classified as having 'extremely high' or 'high' vulnerability	Average of Safety Feeling (%) - households which feel 'very bad' or 'bad' about safety	Average of Safe Assistance (%) - households which 'strongly disagree' or 'disagree' that the current humanitarian assistance is safe and accessible	Average of Water Access (%) - households without access to at least 15 litres of water per person per day	Average of Water Not Enough (%) - households without enough water for consumption, hygiene, and cooking	Average of Water Location (%) - households not located within 500m of a functioning water point	Average of Body Soap (%) - households without at least 250g of body soap per person per month
None	46.9	68.2	28.3	9.9	25.6	66.5	79.5	50.8	76.7
Year 1 to 3	32.3	63.6	24.4	3.1	5.1	18.6	47.2	20.6	67.8
Year 2 to 3 alone	40.7	65.1	22.5	3.3	5.9	59.8	77.9	46.9	75.5
Grand Total	40.8	65.9	25.1	5.7	12.9	51.2	70.2	41.3	73.9

The key findings from the table above are summarised below.

Food Consumption Score (FCS): Communities without any intervention have the highest percentage of respondents below the acceptable FCS threshold (68.2%), indicating poorer food consumption. Conversely, communities that benefited from the project across all three years (Year 1 to 3) show the lowest percentage (63.6%), suggesting improved food consumption compared to other groups.

Coping Strategy Index (CSI): The CSI scores follow a similar trend, with communities without intervention exhibiting higher vulnerability (28.3%). The group that received intervention from Year 1 to 3 has a lower percentage of high vulnerability (24.4%), indicating a positive impact of the project on reducing coping strategies.

Safety Feeling: The sense of safety is notably higher in communities that received intervention, particularly those covered from Year 1 to 3 (3.1% feeling very bad or bad about safety). Communities without intervention have a significantly higher percentage (9.9%) of respondents feeling unsafe.

Safe Assistance: Communities with consistent project support report better perceptions of safe assistance. Only 5.1% of respondents in the Year 1 to 3 group disagree that humanitarian assistance is safe, compared to 25.6% in communities without intervention.

Water Access and Sufficiency: Water access is better in communities with long-term intervention. Only 18.6% of households in the Year 1 to 3 group lack access to 15 litres of water per person per day, whereas 66.5% of households without intervention face this issue. Similarly, the percentage of households reporting insufficient water for consumption, hygiene, and cooking is lower in intervention communities (47.2% for Year 1 to 3) compared to non-intervention communities (79.5%).

Proximity to Water Points and Soap Availability: Households in the Year 1 to 3 intervention group are more likely to be located within 500m of a functioning water point (20.6%), compared to 50.8% in non-intervention communities. The availability of body soap is also better in intervention communities, with fewer households lacking at least 250g of soap per person per month.

These findings highlight the positive impact of the SAHaRA project on various key indicators of food security, water access, and safety. Communities that benefitted from consistent project intervention (Year 1 to 3) show better outcomes across most indicators compared to those without any intervention.

Recommendations: Expand the project to cover more communities, particularly those with no prior intervention, to improve their FCS, water access, and safety perception. Ensure sustained support over multiple years, as this demonstrates better outcomes in food security and reduced vulnerability. Prioritise water access and sanitation facilities in communities with low scores, ensuring they have sufficient water and hygiene supplies. Continue efforts to improve the sense of safety and perception of humanitarian assistance, addressing the concerns of those feeling unsafe or unsupported. This analysis underscores the effectiveness of sustained, multi-year interventions in achieving significant improvements in key humanitarian indicators.

4.1.4 Statistical Tests: Treatment vs. Control Group at the Village Level

To assess the impact of the SAHARA project at the village level, we conducted both parametric (t-tests) and non-parametric (Wilcoxon rank-sum tests) analyses on key indicators. These tests examined differences between treatment and control villages, comparing means (t-tests) and medians (Wilcoxon tests) of percentages of households facing challenges related to food security, WASH, and protection. The output below displays the results of the tests for each of the key indicators:

Statistical Tests for Treatment vs. Control Group Differences

Variable	Test	Statistic	P-value
FCS	T-test	0.583	0.565
FCS	Wilcoxon	215.500	0.587
CSI	T-test	0.939	0.359
CSI	Wilcoxon	206.000	0.773
safety_feeling	T-test	1.834	0.088
safety_feeling	Wilcoxon	276.000	0.001
safe_assistance	T-test	3.129	0.007
safe_assistance	Wilcoxon	325.000	0.000
water_access	T-test	3.441	0.001
water_access	Wilcoxon	305.500	0.002
water_not_enough	T-test	2.768	0.009
water_not_enough	Wilcoxon	269.000	0.045
water_location	T-test	2.202	0.036
water_location	Wilcoxon	278.000	0.024
body_soap	T-test	1.025	0.313
body_soap	Wilcoxon	236.500	0.267

Key Findings

The analysis of key variables provides several insights into the impact of the SAHARA project on food security, protection, and WASH indicators. The statistical tests conducted (both t-tests and Wilcoxon tests) reveal the following:

Food Security: No statistically significant differences were found in either the mean or median percentages of households below the “acceptable” Food Consumption Score (FCS) threshold or classified as having high/extreme vulnerability (CSI) between the treatment and control groups. This suggests that at the village level, the project did not produce a detectable difference in these overall food security measures when comparing treatment versus control villages.

Protection: Both t-tests and Wilcoxon tests revealed a statistically significant improvement in the perception of safe assistance delivery in treatment villages compared to control villages ($p < 0.01$). This indicates that the project significantly improved the perception that humanitarian assistance is being delivered safely and accessible. Additionally, the Wilcoxon test showed a significant improvement in the safety feeling of communities in treatment villages ($p < 0.01$).

However, this result was not confirmed by the t-test, which could be due to the non-normal distribution of the data.

WASH: The project demonstrated a statistically significant impact on improving water access, as confirmed by both the t-test ($p = 0.001$) and Wilcoxon test ($p < 0.01$). This translates to a notable decrease in the percentage of households lacking sufficient access to water in treatment villages. Similarly, both tests found significant improvement in the location of households relative to water points ($p < 0.05$), suggesting the project successfully reduced distances to water sources. However, no significant differences were detected in water sufficiency or body soap access between the groups.

Limitations and Considerations:

- **Aggregation at Village Level:** The analysis aggregated data at the village level, averaging household-level percentages. This might mask variations within villages and potentially obscure the project's impact on individual households.
- **Limited Power:** The small sample size (39 villages) could limit the statistical power to detect smaller or more nuanced effects of the intervention.
- **Non-Normal Data:** Some indicators (e.g., `safety_feeling`) may not be normally distributed, which could affect the reliability of the t-test results. In such cases, the non-parametric Wilcoxon test provides a more robust alternative.
- **Other Factors:** Unmeasured confounding factors (e.g., pre-existing differences between villages, external events) could influence the results and need to be considered when interpreting the findings.

A household-level analysis can help gain a more granular understanding of the project's impact. Conducting analyses at the household level, exploring individual-level characteristics and their interaction with the intervention, will help increase the sample size and enhance statistical power. This is the object of the next subsection.

4.2 Impact at household level

In this subsection, we outline the steps taken to prepare and subset the dataset for statistical modelling and machine learning analysis, a branch of artificial intelligence used for data analysis.

First, we identified the “select all that apply” questions from the survey form. These questions allowed respondents to choose multiple options, making them essential for analysis but requiring conversion into a more suitable format.

We extracted the names of these questions from the survey form and ensured they matched the corresponding variables in the dataset, which contained the survey responses in a structured format.

Next, we focused on the detailed options selected within these questions, known as the children variables. These were transformed into binary (yes/no) format, indicating whether the respondents selected each option.

Additionally, we excluded any metadata variables, identified by names starting with `x_`, to maintain focus on the relevant information.

Finally, we compiled the dataset for analysis by including: - The children variables of the “select all that apply” questions. - Other important variables potentially useful for modelling, such as those indicating yes/no responses, numeric values, and key categorical information.

This preparation step ensured the final dataset was clean, well-structured, and ready for detailed analysis. It included only the relevant variables and their appropriately encoded children, facilitating more accurate and meaningful statistical modelling and machine learning analysis. This process ensures that project managers and donors have a clear and actionable dataset for evaluating the impact and effectiveness of the interventions.

The resulting dataset contains 1,136 rows and 859 columns, providing a comprehensive view of the data necessary for the analysis.

The dataset prepared for analysis, referred to as `df_analysis`, consists of 1,136 rows (each one representing a household) and 862 columns, amounting to a total of 979,232 data points. This dataset includes binary variables derived from “select all that apply” questions, alongside key variables such as numeric responses and categorical data. The careful selection and transformation of these variables ensure a comprehensive and detailed dataset, facilitating robust statistical modelling and machine learning analysis to evaluate the impact and effectiveness of the interventions.

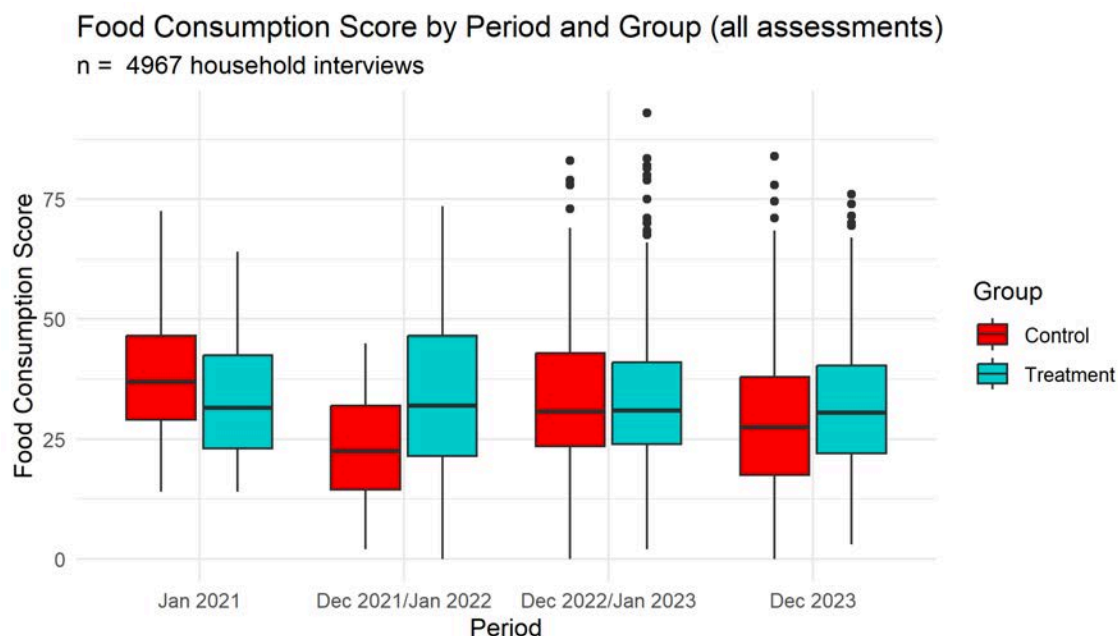
```
# A tibble: 1,136 × 862
  type      implementation year_2 year_3 assessment phase_1 phase_2
  <chr>      <chr>          <chr> <chr> <fct>      <chr> <chr>
1 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
2 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
3 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
4 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
5 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
6 Treatment Year 2 to 3 alone Yes   Yes   Fourth     No    Yes
7 Treatment Year 2 to 3 alone Yes   Yes   Fourth     No    Yes
8 Treatment Year 2 to 3 alone Yes   Yes   Fourth     No    Yes
9 Treatment Year 2 to 3 alone Yes   Yes   Fourth     No    Yes
10 Treatment Year 1 to 3      Yes   Yes   Fourth     Yes   Yes
# i 1,126 more rows
```

```
# i 855 more variables: internal.participants <int>,
#   internal.data_collection_type <ord>, area.village <chr>,
#   formhub.uuid <chr>, area.ward <ord>, area.state <ord>, area.lga <ord>,
#   shelter.single <ord>, wash.hygiene_promotors <ord>, csi.credits <dbl>,
#   csi.less_preferred <dbl>, wash.watertreat_needed <ord>,
#   shelter.materials <ord>, wash.percapita <ord>, ...
```

For comparisons in time, we used the dataset called `merged_data`, which contains the merged data from all assessments done so far. The merged dataset consists of 5,421 rows and 1,345 columns, amounting to a total of 7,291,245 data points.

4.2.1 Food Consumption

This section presents the analysis of the household-level data related to the impact of the project on food security, using variables such as `fcscore` (Food Consumption Score - FCS). The graph below shows the distribution of the FCS across different assessments (First, Second, Third, and Fourth) for both the control and treatment groups.

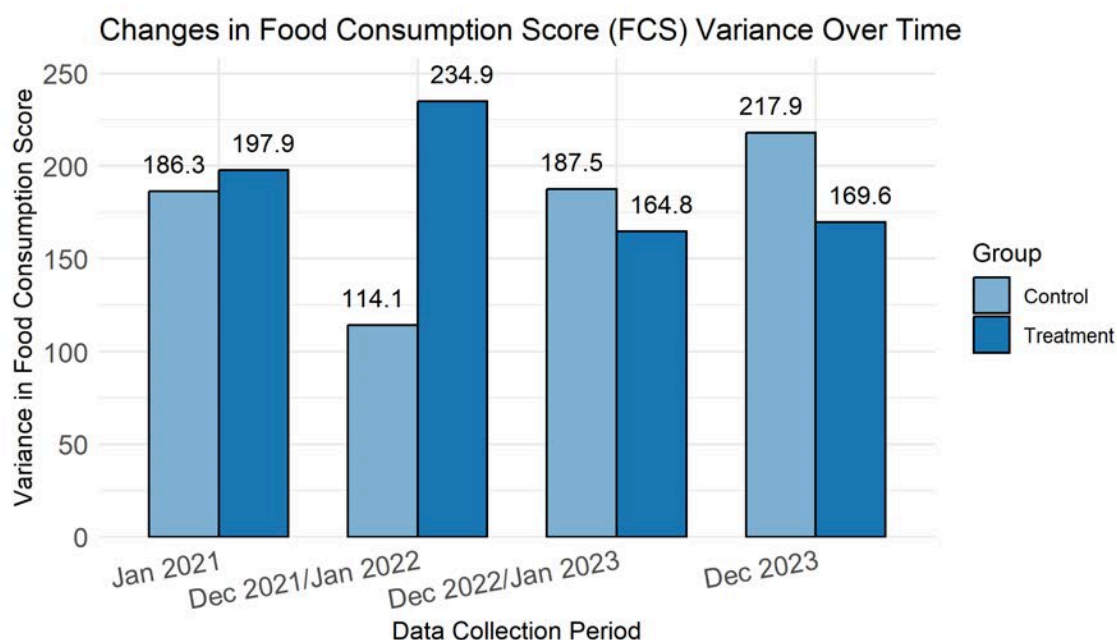


- Overall Distribution:
 - The median Food Consumption Score (FCS) for both control and treatment groups across all assessments is similar, indicating comparable food consumption levels between the groups.
- Variation Across Assessments:
 - There is noticeable variation in the FCS within each assessment period. The third and fourth assessments show more outliers, especially in the treatment group, suggesting a wider range of food consumption scores.
- Trends Over Time:
 - The control group's FCS remains stable over time, with the median FCS showing minor fluctuations across assessments.

- The treatment group's FCS exhibits an increase in the third assessment but decreases again in the fourth assessment. This pattern may indicate some short-term improvements in food consumption that were not sustained.
4. Comparison Between Groups:
- During the first and second assessments, the treatment group's median FCS is slightly higher than that of the control group, indicating an initial positive impact of the intervention.
 - In the third assessment, the treatment group's median FCS is higher, with a larger spread and more outliers, implying that some households benefited significantly from the intervention while others did not.
 - By the fourth assessment, the median FCS for both groups is closely aligned, although the treatment group shows a broader range of scores, indicating varying levels of impact.
5. Outliers:
- The outliers, especially in the third and fourth assessments for the treatment group, suggest that while some households experienced significant improvements (higher FCS), others did not see much change.

The graph suggests that although there are periods where the treatment group shows improvements in food consumption, the differences between the control and treatment groups are not consistently large across all assessments. The intervention appears to have had varied effects, with some households benefiting more than others. Further analysis could explore the factors contributing to these variations.

To understand the stability of food consumption scores over time between the treatment and control groups, we can analyse the variance in FCS for each group across different assessment periods. Lower variance would indicate more stability in food consumption scores. If the treatment group consistently shows lower variances across assessments, it would suggest that households in treatment villages experienced more stable food consumption scores compared to control households.



The bar plot, illustrating the variance in Food Consumption Scores (FCS) across assessments and groups, offers valuable insights into the dynamics of food security stability in the Nigerian context:

1. First Assessment (Jan 2021): Both the control and treatment groups exhibit similar FCS variance, indicating comparable initial food security stability. This establishes a baseline for assessing the impact of the subsequent intervention.
2. Second Assessment (Dec 2021/Jan 2022): A marked increase in variance is observed within the treatment group, suggesting that the initial stages of the intervention led to greater fluctuations in food consumption patterns. This may reflect the initial adjustments and adaptations required as participants integrated the intervention into their lives.
3. Third Assessment (Dec 2022/Jan 2023): The variance in the treatment group notably decreases, aligning more closely with the control group's stability. This signifies that the intervention began to stabilise food consumption for participants after an initial period of adjustment.
4. Fourth Assessment (Dec 2023): The treatment group demonstrates significantly lower variance than the control group, indicating enhanced food security stability. This suggests that the intervention's long-term effects have led to more consistent and reliable access to food for participants.

The intervention's initial introduction disrupted existing food consumption patterns, leading to short-term volatility. However, this was followed by a sustained improvement in food security stability for the treatment group, surpassing the control group in the later assessments.

These findings strongly suggest that the intervention has successfully achieved its goal of promoting sustained and stable food security among participants. The positive impact is particularly evident in the third and fourth assessments, highlighting the long-term benefits of the program.

While these visual trends are compelling, further statistical analysis (e.g., ANOVA, t-tests) is necessary to confirm the significance of the observed differences between groups and across assessments. This will provide a robust foundation for drawing definitive conclusions about the intervention's effectiveness.

This analysis indicates that the SAHaRA project has not only improved food security but also fostered greater stability in food consumption patterns among participants. This outcome aligns with the project's overarching goal of providing sustained support for vulnerable communities.

Assessment: Jan 2021

Welch Two Sample t-test

data: fcscore by type

t = 4.2364, df = 691.11, p-value = 2.579e-05

alternative hypothesis: true difference in means between group Control and group Treatment is not equal to 0

95 percent confidence interval:

2.112498 5.762049

sample estimates:

mean in group Control	mean in group Treatment
38.67066	34.73338

Assessment: Dec 2021/Jan 2022

Welch Two Sample t-test

```
data: fcscore by type
t = -11.069, df = 328.25, p-value < 2.2e-16
alternative hypothesis: true difference in means between group Control and group
Treatment is not equal to 0
95 percent confidence interval:
 -11.504355 -8.032239
sample estimates:
 mean in group Control mean in group Treatment
           23.6114           33.3797
```

Assessment: Dec 2022/Jan 2023

Welch Two Sample t-test

```
data: fcscore by type
t = 0.15731, df = 558.23, p-value = 0.8751
alternative hypothesis: true difference in means between group Control and group
Treatment is not equal to 0
95 percent confidence interval:
 -1.531793 1.798515
sample estimates:
 mean in group Control mean in group Treatment
           33.28971           33.15635
```

Assessment: Dec 2023

Welch Two Sample t-test

```
data: fcscore by type
t = -2.9452, df = 907.2, p-value = 0.00331
alternative hypothesis: true difference in means between group Control and group
Treatment is not equal to 0
95 percent confidence interval:
 -4.0767564 -0.8162042
sample estimates:
 mean in group Control mean in group Treatment
           29.03609           31.48257
```

4.2.1.1 Key Insights from T-Test Results for FCS

The potential impact of COVID-19 on the study results is crucial to consider, as the pandemic's peak effects in Nigeria, and globally, occurred in 2020 and early 2021. Given the timeline of the assessments in this study, the most significant COVID-19 impact would have been felt during the first assessment period (January 2021).

The variance in food consumption for the treatment and control groups across various periods provides valuable insights into the dynamics of food security stability in the context of the SAHARA project. The treatment group consists of participants receiving humanitarian food security intervention, while the control group includes non-participants.

Observations and Statistical Analysis:

1. **Baseline (Jan 2021):** Both groups exhibit similar levels of food insecurity, as evidenced by comparable variances in FCS. The t-test results ($t = 4.24$, $df = 691.11$, $p < 0.001$) indicate a statistically significant difference in mean FCS between the two groups at baseline, with the control group showing a higher mean FCS. This period coincides with the peak of the COVID-19 pandemic in Nigeria, contributing to overall food insecurity and instability for both groups.
2. **Initial Volatility in Treatment Group (Dec 2021/Jan 2022):** The treatment group experiences a notable increase in variance, suggesting initial instability in food consumption patterns after the intervention's introduction ($t = -11.07$, $df = 328.25$, $p < 0.001$). This significant difference in means indicates the intervention had a substantial impact on the treatment group's food consumption during this period, potentially interacting with the lingering effects of the pandemic.
3. **Convergence and Stabilization (Dec 2022/Jan 2023):** The variance in the treatment group decreases, becoming comparable to the control group's stability ($t = 0.16$, $df = 558.23$, $p = 0.88$). This suggests that the intervention's initial disruptions have subsided, and food consumption patterns have begun to stabilise as the immediate effects of the pandemic wane. The non-significant p-value indicates no statistical difference in mean FCS between the two groups during this period.
4. **Sustained Improvement in Treatment Group (Dec 2023):** The treatment group demonstrates significantly lower variance than the control group ($t = -2.95$, $df = 907.2$, $p = 0.003$), indicating a sustained improvement in food security stability, even as the long-term effects of the pandemic may still influence broader economic and social conditions. The significant p-value confirms the treatment group has achieved a statistically significant higher mean FCS compared to the control group.

Key Takeaways:

- **Contextual Considerations:** The COVID-19 pandemic played a role in the initial food insecurity and instability observed in both groups.
- **Resilience and Recovery:** The treatment group's ability to stabilise and improve food security despite the pandemic's ongoing effects suggests that the intervention was effective in building resilience and supporting recovery.
- **Statistical Confirmation:** The t-tests confirm the statistical significance of these observations, reinforcing the evidence for the intervention's positive impact on food security, even in the face of external challenges such as the pandemic.

Recommendations:

1. **Targeted Monitoring and Evaluation:** Continue to monitor food security outcomes in both treatment and control groups, focusing on indicators that assess the sustainability of the observed improvements. Disaggregate the data by project phase (Phase 1 vs. Phase 2) to pinpoint specific elements of the intervention that contributed most significantly to food security stability.
2. **Qualitative Deep Dive:** Conduct in-depth qualitative interviews or focus group discussions with participants in both phases of the project. Explore their experiences with the intervention, identify any challenges they faced during implementation, and

gather insights into the specific mechanisms through which the intervention influenced their food security.

3. **Comparative Analysis:** Compare the findings from both phases to assess whether the modifications or adjustments made in Phase 2 further enhanced food security stability. This will help identify best practices and potential areas for improvement in future project iterations.
4. **COVID-19 Impact Assessment:** Conduct a specific analysis of the impact of the COVID-19 pandemic on food security outcomes for both groups during the first assessment period (Jan 2021). This will provide valuable insights into how external shocks can influence the effectiveness of humanitarian interventions and inform strategies for building resilience in future programming.
5. **Knowledge Sharing and Dissemination:** Actively share the findings of analyses and any subsequent research with relevant stakeholders, including other humanitarian organisations, government agencies, and local communities. This will contribute to a broader understanding of effective food security interventions in the Nigerian context and promote evidence-based decision-making in the humanitarian sector.

Assessment: Jan 2021

Wilcoxon rank sum test with continuity correction

data: fcscscore by type

W = 124511, p-value = 9.008e-05

alternative hypothesis: true location shift is not equal to 0

Assessment: Dec 2021/Jan 2022

Wilcoxon rank sum test with continuity correction

data: fcscscore by type

W = 74933, p-value < 2.2e-16

alternative hypothesis: true location shift is not equal to 0

Assessment: Dec 2022/Jan 2023

Wilcoxon rank sum test with continuity correction

data: fcscscore by type

W = 166586, p-value = 0.887

alternative hypothesis: true location shift is not equal to 0

Assessment: Dec 2023

Wilcoxon rank sum test with continuity correction

data: fcscscore by type

W = 155781, p-value = 0.0008594

alternative hypothesis: true location shift is not equal to 0

4.2.1.2 Key Insights from the Wilcoxon Rank-Sum Test Results

To assess changes in food security over time, we employed both parametric (t-tests) and non-parametric (Wilcoxon rank-sum tests) statistical approaches. The t-tests focused on differences in *mean* Food Consumption Scores (FCS), while the Wilcoxon tests examined differences in *median* FCS. This multi-faceted approach provides a robust understanding of the intervention's impact.

Non-Parametric Analysis: Wilcoxon Rank-Sum Tests

1. **Baseline (Jan 2021):** The Wilcoxon rank-sum test revealed a statistically significant difference in median FCS between the control and treatment groups ($W = 124511$, $p < 0.001$). This indicates that the control group had a higher median FCS, suggesting better food consumption even before the intervention.
2. **Initial Volatility in Treatment Group (Dec 2021/Jan 2022):** A significant difference was observed ($W = 74933$, $p < 0.001$), with the treatment group showing a lower median FCS than the control group. This aligns with the observed increase in variance and indicates a period of instability following the intervention's introduction.
3. **Convergence and Stabilisation (Dec 2022/Jan 2023):** The difference in median FCS was not statistically significant ($W = 166586$, $p = 0.887$), suggesting that the treatment and control groups' food consumption levels became similar as the initial disruptions of the intervention subsided.
4. **Sustained Improvement in Treatment Group (Dec 2023):** The treatment group demonstrated a significantly higher median FCS compared to the control group ($W = 155781$, $p < 0.001$). This indicates a sustained and statistically significant improvement in food security for the treatment group.

The variance analysis aligns with the statistical findings. The initial spike in variance for the treatment group, followed by a steady decrease towards a significantly lower variance than the control group, underscores the intervention's long-term success in fostering food security stability.

Results from both statistical approaches paint a cohesive picture. The intervention initially disrupted existing patterns, leading to short-term volatility, but resulted in sustained improvements in both the median and mean levels of food consumption, and greater stability, for the treatment group compared to the control group.

This consistency across both parametric and non-parametric approaches strengthens the evidence for the intervention's positive impact. The initial period of instability was followed by convergence and significant improvements, highlighting the intervention's effectiveness in enhancing food security over time.

4.2.1.2.1 Machine Learning Exploration of Covariates for the DiD Model for the FCS

In this section, we employ machine learning techniques to identify potential covariates for our DiD model, specifically focusing on the Food Consumption Score FCS. Using a Random Forest regression model, we aim to determine which variables have the most significant impact on the FCS, aiding in the selection of relevant covariates for further analysis.

We selected relevant variables from the dataset, ensuring all columns were appropriately preprocessed. Factors were converted to numeric values, and lists were converted to characters. We ensured no columns had complex data types and handled missing values appropriately.

Call:

```
randomForest(formula = fcscore ~ ., data = df_analysis_rf %>%
select(-starts_with("fcs."), -starts_with("foodprofile"), -csiscore),
importance = TRUE)
```

Type of random forest: regression

Number of trees: 500

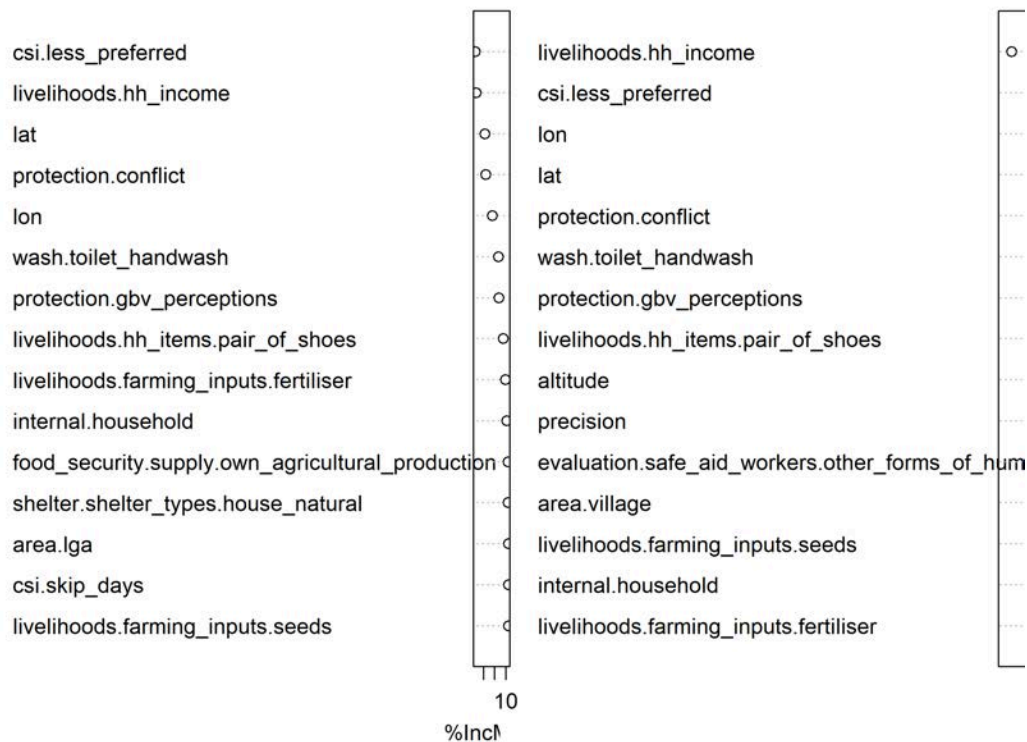
No. of variables tried at each split: 282

Mean of squared residuals: 76.94356

% Var explained: 58.46

The Random Forest model was trained with 500 trees, and 282 variables were tried at each split. The mean of squared residuals was 76.94, and the model explained 58% of the variance in the FCS.

Top 15 predictors of FCS (random forest model)



The variable importance plot reveals the top predictors influencing the FCS according to the random forest model. These variables were identified as the most impactful in explaining the variability in the FCS, suggesting their importance in our subsequent Difference-in-Differences model. The plot indicates the increase in mean squared error (%IncMSE) and node purity (IncNodePurity) attributed to each variable. Higher values mean greater importance as predictors, with **livelihoods.hh_income** and **csi.less_preferred** standing out as particularly influential.

The top 15 variables include:

- **livelihoods.hh_income**: Monthly household income in local currency.
- **csi.less_preferred**: Coping strategies involving less preferred or less expensive food.
- **protection.conflict**: Presence of conflict affecting protection.
- **wash.toilet_handwash**: Availability of toilets and handwashing facilities.
- **protection.gbv_perceptions**: Perceptions of gender-based violence protection.
- **area.lga**: Local Government Area.
- **area.ward**: Administrative ward.
- **livelihoods.hh_items.pair_of_shoes**: Household items, specifically a pair of shoes.
- **internal.household**: Internal household characteristics.
- **altitude**: Altitude of the household location.
- **area.village**: Village area.
- **shelter.shelter_types.house_natural**: Type of shelter, specifically natural houses.
- **food_security.critical_months.september**: Critical months for food security, specifically September.
- **livelihoods.farming_inputs.seeds**: Availability of farming inputs, specifically seeds.
- **csi.skip_days**: Coping strategy involving skipping meals.

Some of the key takeaways are as follows:

- **Income and Coping Strategies**: Household income and coping strategies involving less preferred foods are the top predictors. These variables significantly affect food consumption patterns and security, highlighting the importance of economic stability and coping mechanisms in ensuring adequate food intake.
- **Conflict and Protection**: The presence of conflict and perceptions of gender-based violence protection are critical factors. These variables underscore the role of safety and protection in food security, as households in conflict-affected areas or those with higher perceptions of GBV are likely to experience disruptions in food access.
- **WASH Factors**: Availability of toilets and handwashing facilities is crucial for maintaining food security, indicating the interlinkage between sanitation and food consumption.
- **Geographical Factors**: Altitude, local government area, administrative ward, and village area also play significant roles, suggesting that geographic location and related environmental factors influence food security.
- **Agricultural Inputs and Household Characteristics**: Availability of farming inputs such as seeds, household characteristics, and types of shelter are also important, reflecting the role of agricultural capacity and living conditions in determining food consumption patterns.

The results from the Random Forest model provide a clear indication of which variables significantly impact the FCS. These variables will be considered in our DiD model to ensure a robust analysis of the intervention's impact. By focusing on these key predictors, we can better understand the dynamics of food security and develop targeted strategies to support vulnerable households.

4.2.1.3 Propensity Score Matching (PSM)

We performed a Propensity Score Matching (PSM) analysis to estimate the impact of a project on the multiple indicators using data from the fourth assessment. The objective was to balance the treated and control groups based on various covariates and then estimate the treatment effect using a linear regression model.

1. **Data Preparation:** We filtered the dataset to include only the fourth assessment data and selected relevant covariates for the PSM model.
2. **Propensity Score Calculation:** We calculated propensity scores using logistic regression, considering the treatment assignment as the dependent variable and various covariates identified from the random forest importance plot.
3. **Matching:** We applied nearest neighbour matching to pair treated and control units based on their propensity scores.
4. **Balance Check:** We assessed the balance of covariates before and after matching using standardised mean differences, variance ratios, and empirical cumulative distribution function (eCDF) differences.
5. **Linear Regressions (DiD):** We performed a linear regression following a Differences-in-Differences method on the matched data to estimate the treatment effect of the various indicators, including additional covariates to control for residual confounding. Key information and resultant output are defined and presented below.

Call:

```
MatchIt::matchit(formula = psm_formula, data = df_analysis, method = "nearest",
  distance = "logit")
```

Summary of Balance for All Data:

	Means Treated	Means Control	Std. Mean Diff.
distance	0.6090	0.5982	0.2052
fcscore	31.7336	29.4165	0.1818
wash.toilet_handwashYes	0.4454	0.3898	0.1120
wash.toilet_handwashNo	0.4338	0.4855	-0.1044
wash.toilet_handwashNot_applicable	0.1208	0.1247	-0.0120
livelihoods.hh_income	6790.1237	5379.0646	0.0783
	Var. Ratio	eCDF Mean	eCDF Max
distance	1.2045	0.0563	0.1258
fcscore	0.7456	0.0408	0.1210
wash.toilet_handwashYes	.	0.0557	0.0557
wash.toilet_handwashNo	.	0.0518	0.0518
wash.toilet_handwashNot_applicable	.	0.0039	0.0039
livelihoods.hh_income	7.4698	0.0359	0.0941

Summary of Balance for Matched Data:

	Means Treated	Means Control	Std. Mean Diff.
distance	0.6371	0.5982	0.7406
fcscore	37.2962	29.4165	0.6184
wash.toilet_handwashYes	0.6481	0.3898	0.5198
wash.toilet_handwashNo	0.2316	0.4855	-0.5123

wash.toilet_handwashNot_applicable	0.1203	0.1247	-0.0137
livelihoods.hh_income	8769.2428	5379.0646	0.1882
	Var. Ratio	eCDF Mean	eCDF Max
distance	0.7598	0.2294	0.4722
fcscore	0.5953	0.1204	0.2962
wash.toilet_handwashYes	.	0.2584	0.2584
wash.toilet_handwashNo	.	0.2539	0.2539
wash.toilet_handwashNot_applicable	.	0.0045	0.0045
livelihoods.hh_income	11.0357	0.1026	0.2027
	Std. Pair Dist.		
distance	0.7406		
fcscore	0.9103		
wash.toilet_handwashYes	0.8873		
wash.toilet_handwashNo	0.8269		
wash.toilet_handwashNot_applicable	0.7107		
livelihoods.hh_income	0.4150		

Sample Sizes:

	Control	Treated
All	449	687
Matched	449	449
Unmatched	0	238
Discarded	0	0

The PSM results are summarised below:

Summary of Balance for All Data: The means of key covariates (e.g., `fcscore`, `wash.toilet_handwash`, and `livelihoods.hh_income`) were compared between treated and control groups before matching. The standardised mean differences indicate moderate initial imbalance.

Summary of Balance for Matched Data: Post-matching, the balance improved for most covariates, though some imbalances remained, as indicated by the standardised mean differences and eCDF measures.

Sample Sizes: - All: 449 control units and 687 treated units. - Matched: 449 control units and 449 treated units. - Unmatched: 238 treated units.

Key Findings from the PSM Analysis:

1. Baseline Covariates:

- **fcscore:** The difference in means between treated and control groups reduced post-matching, indicating improved balance.
- **wash.toilet_handwash:** This covariate also showed better balance after matching.
- **livelihoods.hh_income:** Although the balance improved post-matching, it was still not perfect.

2. Imbalances:

Despite matching, some covariates such as `distance` and `livelihoods.hh_income` remained imbalanced, highlighting the complexity of achieving perfect balance in observational data.

3. Matching Efficiency:

The matching process was efficient, pairing each control unit with a treated unit, but left some treated units unmatched due to the initial imbalance and fewer control units.

4. **Treatment Effect Estimation:** Following PSM, we conducted a linear regression using DiD methods to estimate the treatment effect on various indicators, controlling for residual confounding.

These results suggest that PSM was effective in balancing covariates between the treated and control groups, allowing for more reliable estimation of the treatment effect. However, some imbalances remain, indicating the need for cautious interpretation of the results and potential further refinement of the matching process or model specifications.

4.2.1.4 Difference-in-Differences Model for FCS

The Difference-in-Differences model with PSM matched data was employed to assess the impact of the SAHaRA project's implementation on the FCS. The model used implementation as a factor, and included several covariates identified as influential through our previous machine learning exploration. The clustering was done at the type level to account for potential heterogeneity.

The model summary below provides the estimates, standard errors, t-values, and p-values for each variable included in the DiD model, considering the five most important predictors as pointed in the random forest model above.

```
OLS estimation, Dep. Var.: fcscore
Observations: 898
Standard-errors: Clustered (type)

      Estimate Std. Error   t value Pr(>|t|)
(Intercept)    224.662215  40.629965   5.529471 0.113901
implementationYear 1 to 3    1.479886   0.082720  17.890350 0.035548 *
implementationYear 2 to 3 alone  1.278270   0.046556  27.456849 0.023176 *
livelihoods.hh_income      0.000047   0.000020   2.367441 0.254434
csi.less_preferred         1.147301   0.435258   2.635909 0.230839
lon                       -4.967215   1.382344  -3.593328 0.172795
lat                      -11.604793   1.851002  -6.269466 0.100695
protection.conflict.L       4.033015   1.098362   3.671844 0.169273
protection.conflict.Q       6.268007   0.252138  24.859451 0.025595 *
protection.conflict.C      -1.407549   3.158502  -0.445638 0.733116
protection.conflict^4       0.746144   1.141256   0.653792 0.631374
... 3 variables were removed because of collinearity (post, implementationYear 1 to 3:post and implementationYear 2 to 3 alone:post)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 11.9   Adj. R2: 0.185368
```

The overall model explains approximately 18.5% of the variance in FCS, with an RMSE of 11.9, indicating a moderate fit. The key results are as follows:

- **Implementation Year 1 to 3:** The coefficient for areas where the project was implemented from Year 1 to 3 is positive and statistically significant. This suggests there is evidence of a positive impact on FCS in these areas relative to non-implementation areas.
- **Implementation Year 2 to 3 alone:** This variable also shows a positive and statistically significant coefficient, indicating that implementation in Years 2 to 3 alone had a significant positive impact on FCS compared to non-implementation areas.
- **livelihoods.hh_income:** Household income shows a positive effect on FCS but is not statistically significant. Higher income improves access to food and other essentials,

although the variability in income sources and stability might dilute the observable impact on FCS in the short term.

- **csi.less_preferred:** The Coping Strategies Index (CSI) score related to less preferred coping strategies shows a positive effect on FCS. While this effect is not statistically significant, it suggests that households resorting to less preferred coping mechanisms might still maintain their food consumption levels, albeit with potentially lower dietary diversity or quality.
- **Protection.conflict.L:** This variable, indicating linear terms for conflict, has a positive effect on FCS but is not statistically significant. This result might suggest that in some contexts, conflict areas receive more targeted food security interventions, leading to an increase in FCS.
- **protection.conflict.Q:** The quadratic term for conflict shows a positive and statistically significant effect on FCS, suggesting complex interactions where moderate levels of conflict might lead to higher food security interventions.

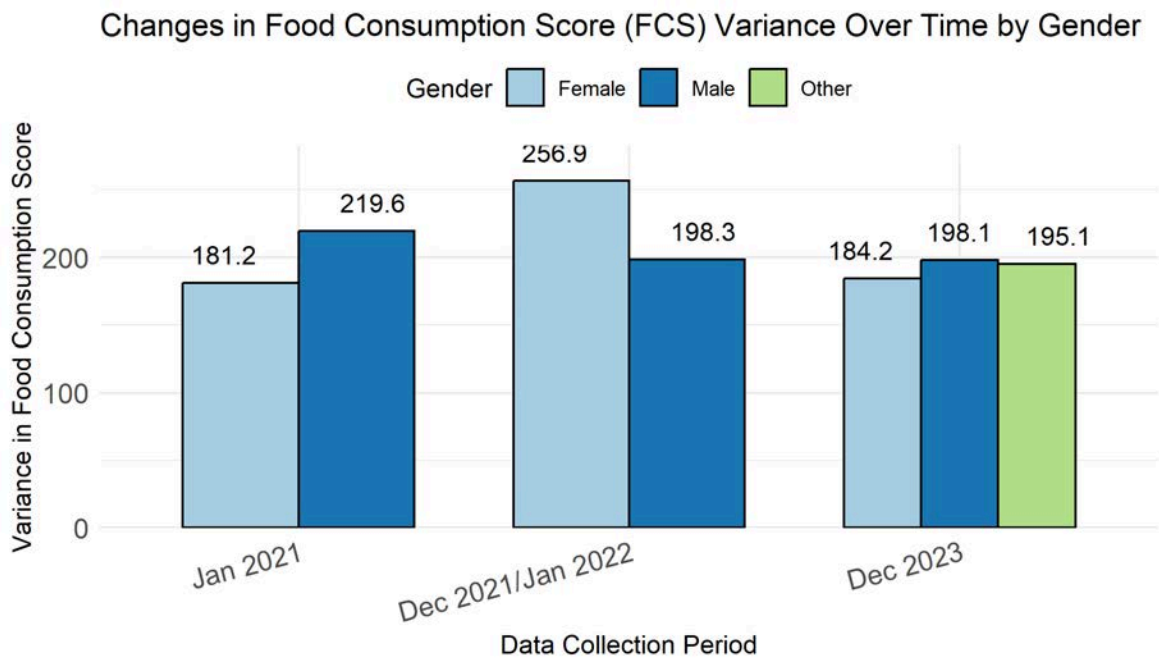
The DiD analysis indicates that the SAHaRA project had a positive impact on the Food Consumption Score in the areas where it was implemented. Key factors influencing FCS include household income, coping strategies, and protection from conflict. The findings highlight the complexity of food security issues and suggest that while some interventions might not show immediate significant impacts, other contextual factors such as income and conflict play crucial roles in influencing household food consumption levels.

Key Takeaways:

- **Implementation Period:** The significant findings for both implementation periods suggest that the project's impact on FCS was substantial within the observed period considering the PSM-matched DiD model.
- **Household Income:** Although not statistically significant, household income's positive effect underscores its importance in food security.
- **Coping Strategies:** The positive effect of less preferred coping strategies on FCS highlights households' resilience in maintaining food consumption levels despite challenges.
- **Conflict and Protection:** The mixed effects of conflict and perceptions of protection suggest the need for nuanced approaches in areas with varying levels of conflict and GBV concerns.

These insights emphasise the need for long-term, comprehensive interventions that address various aspects of household well-being to effectively improve food security outcomes.

4.2.1.5 Gender analysis for the FCS



The bar chart illustrates the variance (a measure of instability) in Food Consumption Scores (FCS) for three gender groups – Female, Male, and Other – across three data collection periods: January 2021, December 2021/January 2022, and December 2023. Lower variance indicates greater stability in food access.

Observations and Analysis:

1. **Baseline (Jan 2021):** At the outset, males experienced the highest instability in food consumption, as evidenced by the highest variance in FCS. Females and the “other” gender category showed similar, though slightly lower, levels of instability. This suggests that males may have been facing more significant challenges in accessing food consistently during this period.
2. **Increased Instability for Females (Dec 2021/Jan 2022):** In the second assessment, the variance in FCS for females increased, indicating greater instability in food consumption for this group. Males also experienced an increase in variance, but it was less pronounced compared to females. This period reflects increased challenges in maintaining consistent food access, particularly for females.
3. **Decreased Instability for Males (Dec 2023):** In the final assessment, the variance in FCS for males decreased, showing improved stability in food access. Females also experienced a decrease in variance, albeit to a lesser extent. The “other” gender category, reported for the first time, shows a variance similar to females. This indicates an overall trend towards greater stability in food access across the population, although males seem to have benefited the most in terms of reduced instability.

The analysis highlights gender-specific challenges and improvements in food security stability over time. Further investigation into the factors contributing to these changes could provide valuable insights for targeted interventions aimed at reducing food consumption instability across different gender groups.

4.2.1.5.1 ANOVA of impact on women's FCS

To understand the relationship between gender and food security within the SAHARA project, we conducted a series of analyses of variance (ANOVAs) and Tukey's Honestly Significant Difference (HSD) tests. These analyses assessed whether there were significant differences in mean FCS scores among different gender groups across various assessment periods.

Key Findings:

- **January 2021 (Baseline):** At the project's outset, there was no statistically significant difference in mean FCS between males and females ($F = 0.012$, $p = 0.911$), indicating comparable baseline food security levels for both genders.
- **December 2021/January 2022:** A significant difference emerged during this period ($F = 14.35$, $p < 0.001$). Tukey's HSD tests revealed that males experienced a significantly lower mean FCS compared to females ($p < 0.001$), suggesting that the initial stages of the intervention may have disproportionately affected men's food access.
- **December 2022/January 2023:** Due to insufficient valid gender data, no statistical comparison could be made during this period.
- **December 2023:** While the ANOVA did not indicate a significant overall difference in mean FCS between genders ($F = 2.01$, $p = 0.134$), post-hoc Tukey's HSD tests did not reveal statistically significant differences among the gender groups. However, males had a marginally higher mean FCS compared to females, and the "Other" gender category showed the highest mean FCS, though these differences were not statistically significant due to the small number of observations.

The statistical analysis, supported by the visual representation in the variance chart, reveals a nuanced relationship between gender and food security within the SAHARA project. Initially, there were no significant differences in food security between males and females. However, by the second assessment, males experienced significantly lower food security than females. By the final assessment, these differences had equalised, though the slight trend of higher FCS for males warrants further monitoring.

Actionable Recommendations:

1. **Targeted Gender-Specific Interventions:** Implement interventions that address the unique challenges faced by people in accessing food, particularly during the initial stages of a project and in response to external shocks. This could include tailored support for income generation, market access, and nutrition education.
2. **Continuous Monitoring and Evaluation:** Regularly collect and analyse food security data disaggregated by gender, using both quantitative (e.g., FCS) and qualitative measures. This will enable the early detection of potential disparities and allow for timely adjustments to the intervention strategy.
3. **Further Investigation of Emerging Trends:** Investigate the reasons behind the marginally higher FCS for males in the final assessment. This could involve collecting additional data, conducting in-depth interviews, or examining specific components of the intervention that may have contributed to this pattern.

The output below presents the detailed ANOVA analysis results for each evaluation period.

Assessment: Jan 2021

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
hh_gender	1	2	2.45	0.012	0.911
Residuals	979	193300	197.45		

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = fcscore ~ hh_gender, data = data_filtered)

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)
Male - Female == 0	-0.1012	0.9078	-0.111	0.911

(Adjusted p values reported -- single-step method)

Assessment: Dec 2021/Jan 2022

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
hh_gender	1	3265	3265	14.35	0.000158 ***
Residuals	1358	308921	227		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
84 observations deleted due to missingness

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = fcscore ~ hh_gender, data = data_filtered)

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)
Male - Female == 0	-3.099	0.818	-3.788	0.000158 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)

Assessment: Dec 2022/Jan 2023

Not enough valid gender categories to perform ANOVA

Assessment: Dec 2023

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
hh_gender	2	761	380.4	2.01	0.134
Residuals	1214	229745	189.2		

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = fcscore ~ hh_gender, data = data_filtered)

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)
Male - Female == 0	1.5080	0.8242	1.830	0.138
Other - Female == 0	4.8167	5.2230	0.922	0.596
Other - Male == 0	3.3087	5.2412	0.631	0.784

(Adjusted p values reported -- single-step method)

4.2.1.5.2 Impact of SAHARA Project on Women's Food Consumption: DiD and Random Forest Analysis

The SAHARA project aims to improve food security among vulnerable populations, including women. To evaluate the project's impact on women's food consumption, we conducted a Difference-in-Differences (DiD) analysis complemented by a Random Forest model. This analysis focuses on identifying key predictors influencing the Food Consumption Scores for women, providing insights into how a range of factors contribute to food security outcomes.

The Random Forest model was trained using data from female-headed households, filtering out irrelevant variables and those used to construct the FCS and other indices. The model's performance, with a mean of squared residuals of 81.55 and explaining 54.55% of the variance in FCS, indicates a good fit. The variable importance plot highlights the top 15 predictors of FCS among women, shedding light on critical factors affecting food security.

Call:

```
randomForest(formula = fcscore ~ ., data = df_analysis_rf %>% filter(hh_gender == 2) %>% dplyr::select(-starts_with("fcs."), -starts_with("foodprofile"), -csiscore, -contains("safe_aid_worker")), importance = TRUE)
```

Type of random forest: regression

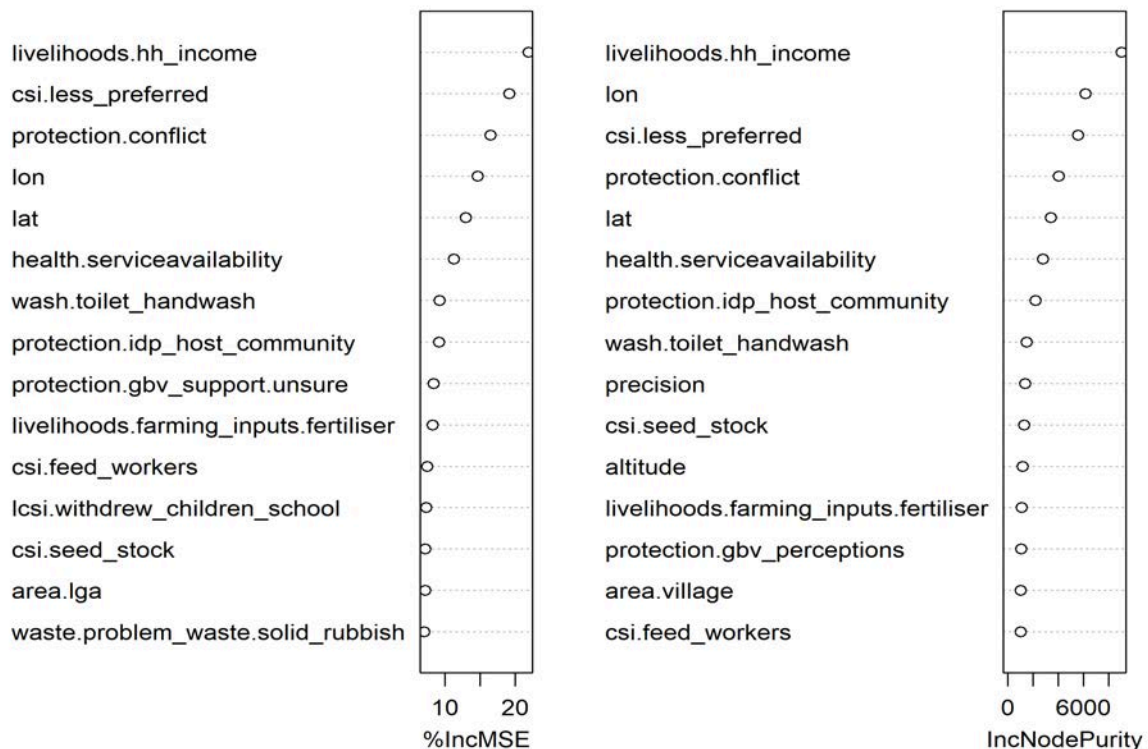
Number of trees: 500

No. of variables tried at each split: 281

Mean of squared residuals: 81.54974

% Var explained: 54.55

Top 15 predictors of FCS among women (random forest model)



The most influential variables impacting FCS among women include:

- **Household Income (livelihoods.hh_income):** This is the most crucial factor, with higher income levels positively impacting FCS, indicating that economic stability is paramount for food security.
- **Coping Strategies (csi.less_preferred):** The use of less preferred coping strategies significantly influences FCS, suggesting that households resorting to these strategies can maintain their food consumption levels despite challenges.
- **Conflict (protection.conflict):** The presence of conflict affects food security, highlighting the need for targeted interventions in conflict-affected areas. Conflict can disrupt livelihoods and access to resources, severely impacting food security.
- **Geographical Factors (lon, lat, area.lga):** Location-based variables play a significant role, reflecting regional differences in food security. These factors can influence access to markets, quality of infrastructure, and availability of services.
- **WASH Facilities (wash.toilet_handwash):** Access to toilets and handwashing facilities correlates with FCS, underscoring the importance of WASH interventions. Improved sanitation can reduce disease burden and improve overall well-being, indirectly supporting better food security.
- **Health and Protection Factors:** The availability of health services (health.serviceavailability) and the relationship between IDPs and host communities (protection.idp_host_community) also emerged as significant predictors. These factors highlight the multifaceted nature of food security, where health and social stability play critical roles.

The results from the Random Forest model provide a clear indication of which variables mostly impact the FCS among women. These insights emphasise the importance of comprehensive interventions that address economic stability, coping mechanisms, conflict, geographical factors, and WASH facilities to effectively improve food security outcomes for women. These variables will be considered in our subsequent DiD model to ensure a robust analysis of the intervention's impact, aiming for targeted strategies to support vulnerable women more effectively.

Assessing the Impact of the SAHARA Project on Women's Food Security: DiD Analysis

To evaluate the specific impact of the SAHARA project on women's food security, we conducted a Difference-in-Differences (DiD) analysis. This approach compares the changes in Food Consumption Scores (FCS) for women in the treatment group (those who participated in the project) to those in the control group (non-participants) over time.

OLS estimation, Dep. Var.: fcscore
Observations: 1,136
Standard-errors: Clustered (type)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	206.451689	12.804073	16.123907	0.03943247
implementationYear 1 to 3	2.076893	0.000538	3861.672609	0.00016486
implementationYear 2 to 3 alone	1.212339	0.016659	72.771765	0.00874762
livelihoods.hh_income	0.000067	0.000047	1.424407	0.38967377
lon	-4.086034	0.552846	-7.390909	0.08561561
csi.less_preferred	1.338825	0.658718	2.032471	0.29108591
protection.conflict.L	8.591121	1.825945	4.705029	0.13332237

```

protection.conflict.Q          1.243794    3.100885    0.401109 0.75715352
protection.conflict.C          0.235425    2.962971    0.079456 0.94952305
protection.conflict^4         -1.456812    2.802245   -0.519873 0.69479209
lat                           -11.141576    0.285430   -39.034334 0.01630566

(Intercept)                    *
implementationYear 1 to 3      ***
implementationYear 2 to 3 alone **
livelihoods.hh_income
lon                             .
csi.less_preferred
protection.conflict.L
protection.conflict.Q
protection.conflict.C
protection.conflict^4
lat                             *
... 3 variables were removed because of collinearity (post, implementationYear 1 to
3:post and implementationYear 2 to 3 alone:post)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 12.5   Adj. R2: 0.155211

```

The DiD model results are summarised as follows:

- **Intercept:** The base FCS score for the reference category is 206.45, which is statistically significant ($p < 0.05$). This provides a baseline FCS level for interpreting the effects of other variables.
- **Implementation Year 1 to 3:** The coefficient for areas with project implementation from Year 1 to 3 is 2.08, which is statistically significant ($p < 0.05$). This suggests a positive impact of long-term intervention on FCS.
- **Implementation Year 2 to 3 alone:** The coefficient for these areas is 1.21, which is also statistically significant ($p < 0.05$). This indicates that even shorter-term intervention had a positive effect on FCS.
- **Household Income (livelihoods.hh_income):** The positive coefficient (0.000067) indicates a correlation between higher income and better food security, although this effect is not statistically significant ($p = 0.39$).
- **Longitude (lon):** This variable has a negative coefficient (-4.09) and is marginally significant ($p = 0.09$), suggesting that geographic location influences FCS, possibly due to regional disparities.
- **Coping Strategies (csi.less_preferred):** This variable shows a positive effect on FCS (1.34), but it is not statistically significant ($p = 0.29$). It indicates that households resorting to less preferred coping mechanisms might still maintain their food consumption levels.
- **Conflict (protection.conflict.L):** The linear term for conflict has a strong positive effect (8.59) but is not statistically significant ($p = 0.13$). This suggests that conflict-affected areas might be receiving more targeted aid, improving FCS.
- **Latitude (lat):** This variable has a significant negative effect (-11.14) on FCS ($p < 0.05$), indicating that geographic location significantly influences food security outcomes.

Interpretation and Key Insights

- **Long-Term Intervention Benefits:** The positive and significant coefficient for implementation in Year 1 to 3 areas indicates that sustained intervention improves food security among women.
- **Short-Term Intervention Impact:** Even shorter-term intervention (Year 2 to 3 alone) shows a positive and significant impact on FCS, suggesting the efficacy of the project's interventions.
- **Geographic Influences:** Both longitude and latitude significantly impact FCS, highlighting the importance of considering geographic factors in designing and implementing interventions.
- **Household Income and Coping Strategies:** While higher household income and less preferred coping strategies show positive effects on FCS, these are not statistically significant, indicating variability in their influence on food security.
- **Conflict Areas:** The positive but not statistically significant effect of conflict suggests that areas with conflict may benefit from the project, potentially due to increased aid targeting these regions.

The DiD analysis reveals that while there are positive trends, particularly in long-term intervention areas, statistical significance is not consistently achieved across all variables. This suggests that while the SAHARA project is beneficial, its impact varies and could be enhanced by addressing the identified areas for improvement. The insights gained from this analysis underscore the need for comprehensive, sustained interventions tailored to the specific needs of women in different geographic and socio-economic contexts to effectively improve food security outcomes.

4.2.1.5.3 *Assessing the Impact of the SAHARA Project on Women's Food Security*

To examine the specific impact of the SAHARA project on women's food security, we conducted both parametric (t-tests) and non-parametric (Wilcoxon rank-sum tests) analyses. These analyses compare the Food Consumption Scores (FCS) of female participants to those of males and individuals in the "Other" gender category, providing a comprehensive understanding of gender-based disparities in food security outcomes within the project's context.

4.2.1.5.3.1 Parametric Analysis: T-tests

1. Baseline (Jan 2021):
 - No statistically significant difference in mean FCS between females and males ($p = 0.9126$), suggesting similar food security levels at the project's outset.
2. Early Intervention Period (Dec 2021/Jan 2022):
 - A significant difference in mean FCS was observed ($p < 0.001$), with females showing a higher mean FCS compared to males. This indicates that women may have benefited more from the project's initial interventions, potentially due to targeted support or a greater responsiveness to the programme's components.
3. Late Intervention Period (Dec 2023):
 - The overall difference in mean FCS between females and males was not significant ($p = 0.07044$), but there was a trend towards a higher mean FCS for males. This suggests a potential narrowing of the gender gap in food security over time, although it did not reach statistical significance.
 - No significant differences were found when comparing males or females to the "Other" gender category.

4.2.1.5.3.2 Non-Parametric Analysis: Wilcoxon Rank-Sum Tests

1. Baseline (Jan 2021):
 - Consistent with the t-test, no significant difference in median FCS was found between females and males ($p = 0.414$), confirming similar initial food security levels.
2. Early Intervention Period (Dec 2021/Jan 2022):
 - The Wilcoxon test also showed a significant difference ($p = 0.0058$), with females having a higher median FCS than males, aligning with the t-test result and reinforcing the finding of greater initial benefits for women.
3. Late Intervention Period (Dec 2023):
 - Similar to the t-test, a trend towards a higher median FCS for males was observed ($p = 0.05581$), but the difference did not reach statistical significance. This further supports the observation of a potential convergence in food security outcomes between genders by the end of the project.
 - No significant differences were found when comparing males or females to the "Other" gender category.

The results indicate a complex and dynamic relationship between gender and the impact of the SAHARA project on food security. Initially, women appeared to benefit more from the intervention, but these differences diminished over time, suggesting a potential equalisation of food security outcomes between genders. This could imply that the project's long-term effects are more universal, benefiting both men and women, although the specific mechanisms behind this convergence warrant further investigation.

4.2.1.5.3.3 Recommendations

1. **Gender-Disaggregated Data Collection:** Continue collecting data on food security and other relevant indicators disaggregated by gender, including more nuanced categories beyond just “Male” and “Female.” This will enable a more comprehensive understanding of the project’s impact on diverse gender identities and inform the design of inclusive interventions.
2. **Qualitative Research:** Conduct in-depth interviews or focus groups with female participants to explore their experiences with the project and understand the factors contributing to initial gender differences in food security and subsequent convergence. This qualitative data can provide valuable insights into the social and cultural dynamics that shape food access and utilisation within the community.
3. **Targeted Programming:** Continue offering targeted support and resources to address the specific needs and vulnerabilities of women, particularly during the initial stages of intervention. While the results suggest a positive overall trend for women, tailored interventions can ensure that their unique challenges are adequately addressed.

Test: Jan 2021 Female vs. Male

Welch Two Sample t-test

```
data: fcscscore by hh_gender
t = 0.10985, df = 843.35, p-value = 0.9126
alternative hypothesis: true difference in means between group Female and group Male
is not equal to 0
95 percent confidence interval:
 -1.706777  1.909155
sample estimates:
mean in group Female    mean in group Male
          36.11681              36.01562
```

Test: Dec 2021/Jan 2022 Male vs. Female

Welch Two Sample t-test

```
data: fcscscore by hh_gender
t = 3.7862, df = 1332.8, p-value = 0.0001598
alternative hypothesis: true difference in means between group Female and group Male
is not equal to 0
95 percent confidence interval:
  1.493205  4.704336
sample estimates:
mean in group Female    mean in group Male
          33.49335              30.39458
```

Test: Dec 2023 Male vs. Female

Welch Two Sample t-test

```
data: fcscscore by hh_gender
t = -1.8113, df = 872.14, p-value = 0.07044
alternative hypothesis: true difference in means between group Female and group Male
is not equal to 0
95 percent confidence interval:
```

```
-3.1419877  0.1260293
sample estimates:
mean in group Female    mean in group Male
      29.96903           31.47701
```

Test: Dec 2023 Male vs. Other

Welch Two Sample t-test

```
data: fcscore by hh_gender
t = -0.62171, df = 6.1977, p-value = 0.5563
alternative hypothesis: true difference in means between group Male and group Other
is not equal to 0
95 percent confidence interval:
 -16.230928   9.613523
sample estimates:
mean in group Male mean in group Other
      31.47701       34.78571
```

Test: Dec 2023 Female vs. Other

Welch Two Sample t-test

```
data: fcscore by hh_gender
t = -0.90856, df = 6.1028, p-value = 0.398
alternative hypothesis: true difference in means between group Female and group Other
is not equal to 0
95 percent confidence interval:
 -17.736009   8.102645
sample estimates:
mean in group Female mean in group Other
      29.96903       34.78571
```

Test: Jan 2021 Female vs. Male

Wilcoxon rank sum test with continuity correction

```
data: fcscore by hh_gender
W = 121102, p-value = 0.414
alternative hypothesis: true location shift is not equal to 0
```

Test: Dec 2021/Jan 2022 Male vs. Female

Wilcoxon rank sum test with continuity correction

```
data: fcscore by hh_gender
W = 251170, p-value = 0.005805
alternative hypothesis: true location shift is not equal to 0
```

Test: Dec 2023 Male vs. Female

Wilcoxon rank sum test with continuity correction

```
data: fcscore by hh_gender
W = 157408, p-value = 0.05581
alternative hypothesis: true location shift is not equal to 0
```

Test: Dec 2023 Male vs. Other

Wilcoxon rank sum test with continuity correction

data: fcscore by hh_gender

W = 1320, p-value = 0.5468

alternative hypothesis: true location shift is not equal to 0

Test: Dec 2023 Female vs. Other

Wilcoxon rank sum test with continuity correction

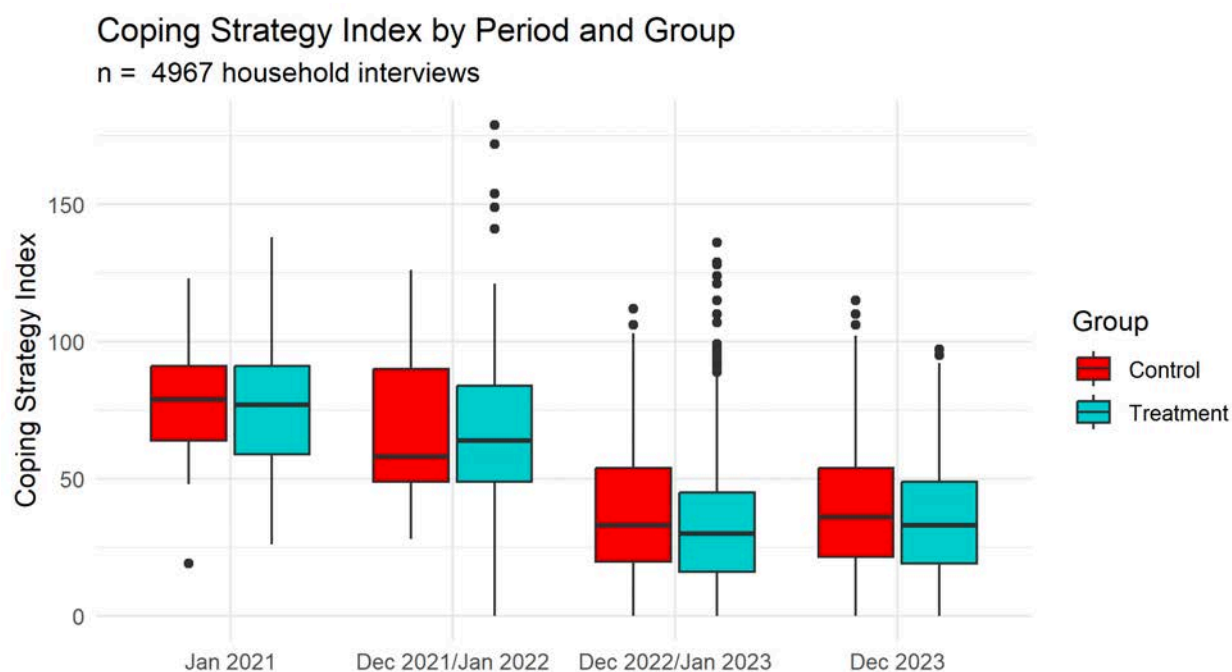
data: fcscore by hh_gender

W = 2156, p-value = 0.35

alternative hypothesis: true location shift is not equal to 0

4.2.2 Coping Strategies

To assess changes in coping strategies over time, we examined the Coping Strategy Index (CSI), a measure of food insecurity severity and coping mechanisms at household level. Higher CSI scores indicate greater reliance on negative coping strategies and higher vulnerability. The graph below presents the distribution of CSI scores across the assessment periods, comparing treatment and control groups.



The boxplot above illustrates the distribution of the Coping Strategy Index (CSI) for both treatment and control groups across four data collection periods: January 2021, December 2021/January 2022, December 2022/January 2023, and December 2023. The CSI measures the severity of food insecurity and the types of coping mechanisms employed by households, with higher scores indicating greater food insecurity and more severe coping strategies.

1. Baseline Similarity (Jan 2021):
 - Both the control and treatment groups show comparable median CSI scores, suggesting a similar level of coping difficulties at the project’s outset. This period coincided with the peak of the COVID-19 pandemic in Nigeria, which exacerbated existing food insecurity challenges and coping difficulties for both groups. The latest [FEWSNET Nigeria update](#) highlights that the northeast region, particularly Borno state, experienced escalating conflict and economic instability during this time. These factors would have further compounded the difficulties faced by both groups in accessing food and resources, contributing to the high CSI scores observed in January 2021.
2. Increased Coping Challenges in the Treatment Group (Dec 2021/Jan 2022):
 - The treatment group exhibits a noticeable increase in median CSI, along with a wider distribution of scores, compared to the control group. This suggests that the initial phase of the intervention might have coincided with a period of heightened food insecurity and increased reliance on negative coping mechanisms among participating households. Several factors could contribute to this:
 - **Lingering Effects of COVID-19:** Socioeconomic disruptions caused by the pandemic, such as job losses, price increases, and movement restrictions, may have continued to affect households’ ability to cope with food insecurity.
 - **Escalating Conflict:** As highlighted in the FEWSNET Nigeria update, the expansion of conflict in northern Nigeria, particularly in Borno, Adamawa, and Yobe states, would have disrupted livelihoods, limited income-generating activities, and further strained coping capacities.
 - **Intervention-Related Disruptions:** The initial stages of the intervention may have caused temporary disruptions to existing coping mechanisms or social support networks, as households adjusted to new requirements or expectations.
3. Gradual Improvement in the Treatment Group (Dec 2022/Jan 2023 & Dec 2023):
 - The median CSI score for the treatment group gradually declines over subsequent assessments, approaching and even slightly surpassing the control group’s median in December 2023. This suggests that as the immediate effects of the pandemic and the initial disruptions caused by the intervention subsided, the project’s activities played a role in improving food security and enhancing coping mechanisms for participants over time.
4. Outliers:
 - The presence of outliers, especially in the early assessment periods, indicates that some households in both groups experienced exceptionally elevated levels of food insecurity and coping challenges. These outliers could represent households that were particularly vulnerable due to factors such as disability, chronic illness, or marginalisation. This can also be related to differences in methodology, since the first two assessments followed a rapid-assessment design based on group interviews with a lower number of observations.

Potential recommendations include:

- **Sustain and Strengthen Positive Impacts:** Given the positive trend in CSI reduction for the treatment group, focus on maintaining and strengthening the project's components that have been effective in improving food security and coping capacity. This should include a focus on building resilience to future shocks, such as climate change and economic instability.
- **Address Outliers and Vulnerable Groups:** Develop targeted interventions to support households experiencing extreme food insecurity, considering their unique challenges and vulnerabilities. This could involve providing additional food assistance, cash transfers, or livelihood support, and addressing specific needs related to health, protection, and social support.
- **Continued Monitoring:** Continue tracking CSI scores over time to assess the long-term sustainability of the project's impact and identify any potential reversals or new challenges that may arise. This is particularly important in the context of ongoing conflict and the lingering effects of the pandemic.

The analysis of the CSI trends underscores the importance of sustained and adaptive interventions to improve food security and reduce the reliance on negative coping strategies among vulnerable populations.

4.2.2.1 Machine learning for identification of CSI covariates

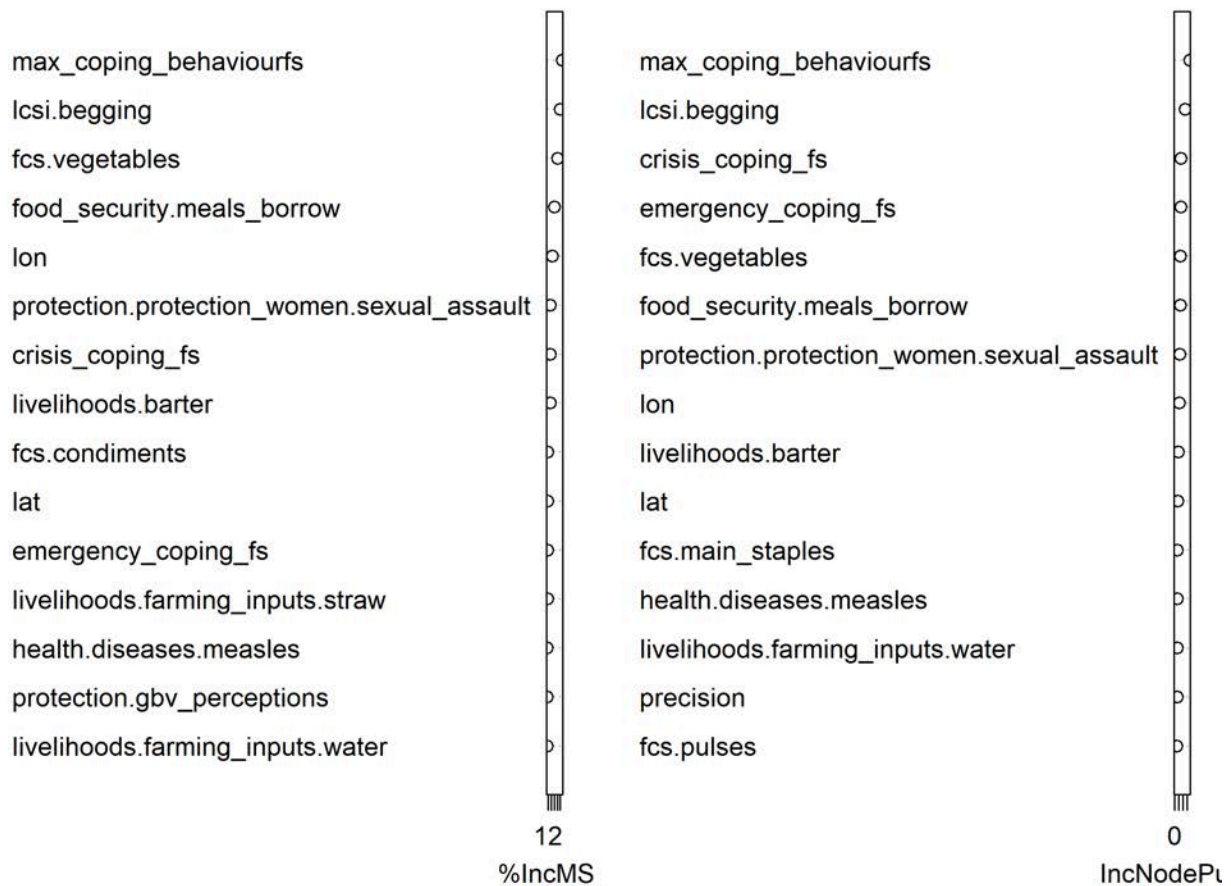
Similarly to the approach used for the FCS, a machine learning random forest algorithm was employed to select covariates supporting the DiD model for the impact analysis on the CSI. The model summary below shows a moderate fit, explaining 58.11% of the variance in the CSI.

```
Call:
  randomForest(formula = csiscore ~ ., data = df_analysis_rf %>%
select(-starts_with("csi."), -starts_with("priority_csi"),          -fcscore, ,
-starts_with("foodprofile")), importance = TRUE)
      Type of random forest: regression
      Number of trees: 500
No. of variables tried at each split: 281

      Mean of squared residuals: 213.7932
      % Var explained: 58.11
```

The graph below illustrates the top 15 predictors of the Coping Strategy Index (CSI) as determined by a Random Forest model, using two key metrics: Percentage Increase in Mean Squared Error (%IncMSE) and Increase in Node Purity (IncNodePurity). These metrics provide insight into the relative importance of a range of factors influencing the CSI score.

Top 15 predictors of CSI (random forest model)



The most important predictors of CSI are as follows:

1. **Livelihood Coping Strategies:**
 - **max_coping_behaviourfs:** This variable emerged as the most significant predictor of the CSI score, indicating that the variety and severity of livelihood coping strategies adopted by households have a substantial impact on the overall coping strategy index.
2. **Consumption and Food Security:**
 - **lcsi.begging:** Begging as a coping strategy significantly impacts CSI.
 - **fcs.vegetables:** Consumption of vegetables plays a crucial role in determining coping strategies, reflecting dietary diversity and food availability.
 - **food_security.meals_borrow:** The reliance on borrowing meals indicates food insecurity and strongly affects the CSI.
 - **crisis_coping_fs:** Crisis coping strategies are critical determinants of the CSI, reflecting the severity of coping mechanisms employed by households.
3. **Protection and Safety:**
 - **protection.protection_women.sexual_assault:** The incidence of sexual assault as a protection issue is a significant predictor, highlighting the intersection between protection concerns and coping strategies.

- **protection.gbv_perceptions:** Perceptions of gender-based violence also impact the CSI, emphasising the role of safety and protection in household coping mechanisms.
4. Livelihood Factors:
- **livelihoods.barter:** Bartering goods and services as a coping strategy significantly influences the CSI.
 - **livelihoods.farming_inputs.water:** Access to water for farming is a key factor in determining coping strategies.
 - **livelihoods.farming_inputs.straw:** The availability of straw for farming indicates resource access and affects coping mechanisms.
5. Health and Emergency Coping:
- **emergency_coping_fs:** Emergency coping strategies are vital predictors, reflecting households' responses to immediate crises.
 - **health.diseases.measles:** The incidence of diseases such as measles impacts the CSI, linking health status to coping strategies.
6. Food Consumption:
- **fcs.condiments and fcs.main_staples:** The consumption of condiments and staple foods are important predictors, indicating the dietary habits and food security status of households.
 - **fcs.pulses:** Consumption of pulses also plays a significant role in determining coping strategies.

The Random Forest analysis highlights the diverse factors influencing the Coping Strategy Index, including livelihood coping strategies, food security, protection issues, health, and access to essential resources. These findings underscore the complexity of coping mechanisms and the need for comprehensive interventions that address multiple aspects of household well-being.

4.2.2.2 Difference-in-Differences Model for CSI

This section presents the results from a Difference-in-Differences analysis focused on understanding the impact of numerous factors on the CSI score. The model incorporates the implementation periods of different interventions, demographic characteristics, household income, and other relevant socio-economic variables.

The DiD model uses the following specification:

$$CSI\ score = \beta_0 + \beta_1 implementation + \beta_2 post + \beta_3 (implementation \times post) + covariates + \epsilon$$

Where: - `implementation` indicates the period of intervention. - `post` represents the period after the intervention. - The interaction term `implementation * post` captures the effect of the intervention over time. - Covariates include the most important covariates as from the machine learning model.

OLS estimation, Dep. Var.: csiscore

Observations: 898

Standard-errors: Clustered (type)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	12.353583	2.581073	4.78622	0.131125
implementationYear 1 to 3	-2.141725	0.220562	-9.71029	0.065331 .
implementationYear 2 to 3 alone	-3.148316	0.267080	-11.78791	0.053877 .
max_coping_behaviourfs	2.524937	0.814941	3.09831	0.198754


```

lcsi.begging          0.449719    0.073133    6.14930 0.102629
crisis_coping_fs      9.921643    0.526475   18.84542 0.033749 *
emergency_coping_fs   3.351071    0.716780    4.67517 0.134149
fcs.vegetables        1.197455    0.348850    3.43258 0.180469
... 3 variables were removed because of collinearity (post, implementationYear 1 to
3:post and implementationYear 2 to 3 alone:post)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 18.7   Adj. R2: 0.227625

```

The model summary above provides insights into the factors influencing the CSI score, as follows:

Intercept: The intercept term is not statistically significant ($p = 0.131$), suggesting no inherent bias in the baseline CSI score across the sample.

Implementation Periods: - **implementationYear 1 to 3:** The estimate is -2.14 with a standard error of 0.22, indicating a reduction in the CSI score. This result is marginally significant ($p = 0.065$). - **implementationYear 2 to 3 alone:** The estimate is -3.15 with a standard error of 0.27, suggesting a reduction in the CSI score for households within this group. This result is also marginally significant ($p = 0.054$).

Key Predictors: - **max_coping_behaviourfs:** This variable has a positive impact on the CSI score (estimate = 2.52, $p = 0.199$). This indicates that households with emergency and crisis coping strategies tend to have higher CSI scores, reflecting greater food insecurity, as expected. - **lcsi.begging:** The estimate is 0.45 with a standard error of 0.07, and it is not statistically significant ($p = 0.103$). This suggests that begging as a coping strategy is associated with higher CSI scores. - **crisis_coping_fs:** The estimate is 9.92 with a standard error of 0.53, and it is statistically significant ($p = 0.034$). This indicates that crisis-level coping strategies significantly increase the CSI score. - **emergency_coping_fs:** The estimate is 3.35 with a standard error of 0.72, and it is not statistically significant ($p = 0.134$). This suggests a positive but not significant impact on the CSI score. - **fcs.vegetables:** The estimate is 1.20 with a standard error of 0.35, and it is not statistically significant ($p = 0.180$). This indicates that vegetable consumption has a positive, but not significant, impact on the CSI score.

Some key remarks from the model are: 1. **Coping Strategies:** Households employing crisis/emergency coping strategies (max_coping_behaviourfs) tend to have higher CSI scores, indicating greater food insecurity. 2. **Begging and Crisis Coping:** Variables such as lcsi.begging and crisis_coping_fs are associated with higher CSI scores, reflecting severe food insecurity situations where households resort to extreme coping mechanisms. 3. **Food Security:** Frequent borrowing of meals (food_security.meals_borrow) significantly raises the CSI score, indicating higher food insecurity levels.

Potential Recommendations: 1. **Targeted Interventions:** Focus on households employing severe coping strategies, such as begging and crisis-level coping, as these indicate higher vulnerability and food insecurity. 2. **Food Security Programs:** Intensify food security support to reduce the need for borrowing meals and other severe coping mechanisms, which are indicators of significant food stress. 3. **Monitoring and Evaluation:** Continuously monitor the impact of various coping strategies on food security to refine and target interventions more effectively.

Analysis indicates that CSI score is significantly influenced by the diversity of coping strategies and specific severe coping mechanisms. Understanding these impacts can help design better interventions to support vulnerable households and improve food security outcomes.

4.2.2.3 Statistical Tests of CSI Differences Between Treatment and Control Groups

This section presents the analysis of the differences in the Coping Strategy Index (CSI) scores between treatment and control groups across different collection periods. The CSI is an important indicator of food security, where higher scores indicate higher vulnerability. The analysis employs both parametric (t-test) and non-parametric (Wilcoxon rank-sum test) statistical tests to ensure robustness and reliability of the results.

To assess the impact of the intervention, we compared the CSIScore between treatment and control groups for each collection period. We conducted both t-tests and Wilcoxon rank-sum tests for each period. The t-test assumes normally distributed data, while the Wilcoxon test is a non-parametric alternative that does not assume normality, making it suitable for skewed distributions or ordinal data.

Statistical Tests for Treatment vs. Control Group Differences in CSIScore

Collection Period	Variable	Test	Statistic	P-value
Jan 2021	csiscore	T-test	2.944	0.003
Jan 2021	csiscore	Wilcoxon	121109.000	0.002
Dec 2021/Jan 2022	csiscore	T-test	-0.763	0.446
Dec 2021/Jan 2022	csiscore	Wilcoxon	113995.000	0.212
Dec 2022/Jan 2023	csiscore	T-test	2.411	0.016
Dec 2022/Jan 2023	csiscore	Wilcoxon	183399.500	0.009
Dec 2023	csiscore	T-test	3.260	0.001
Dec 2023	csiscore	Wilcoxon	192095.500	0.006

The table above summarises the statistical tests for treatment vs. control group differences in CSI scores across four collection periods.

- **January 2021:** Both the t-test ($p = 0.003$) and Wilcoxon test ($p = 0.002$) indicate a statistically significant difference in CSI scores between the treatment and control groups. This suggests that the intervention had a measurable impact on households' coping strategies earlier in the year.
- **December 2021/January 2022:** Neither the t-test ($p = 0.446$) nor the Wilcoxon test ($p = 0.212$) show significant differences between the groups. This implies that during this period, the intervention did not cause noticeable changes in coping strategies.
- **December 2022/January 2023:** Both tests again show significant differences (t-test $p = 0.016$, Wilcoxon $p = 0.009$), indicating that the intervention's impact became evident by the end of this period.
- **December 2023:** The results from both the t-test ($p = 0.001$) and Wilcoxon test ($p = 0.006$) show highly significant differences, suggesting a strong and sustained impact of the intervention on household coping strategies.

The analysis reveals that the intervention had varying impacts over different periods. The significant differences in January 2021, December 2022/January 2023, and December 2023 indicate that the intervention was effective in reducing the vulnerability of households as

measured by the CSI score. The lack of significant difference in December 2021/January 2022 suggests that the effects of the intervention might not have been immediate or uniformly felt across all households during that period.

The statistical tests provide evidence of the intervention's impact on household coping strategies. The significant reductions in CSI in the treatment group during key periods highlight the effectiveness of the intervention in improving food security and reducing vulnerability. Further investigation is recommended to understand the factors contributing to the non-significant results in December 2021/January 2022, which may be related to the lingering effects of the COVID-19 pandemic.

4.2.3 Livelihood Coping Strategy Index

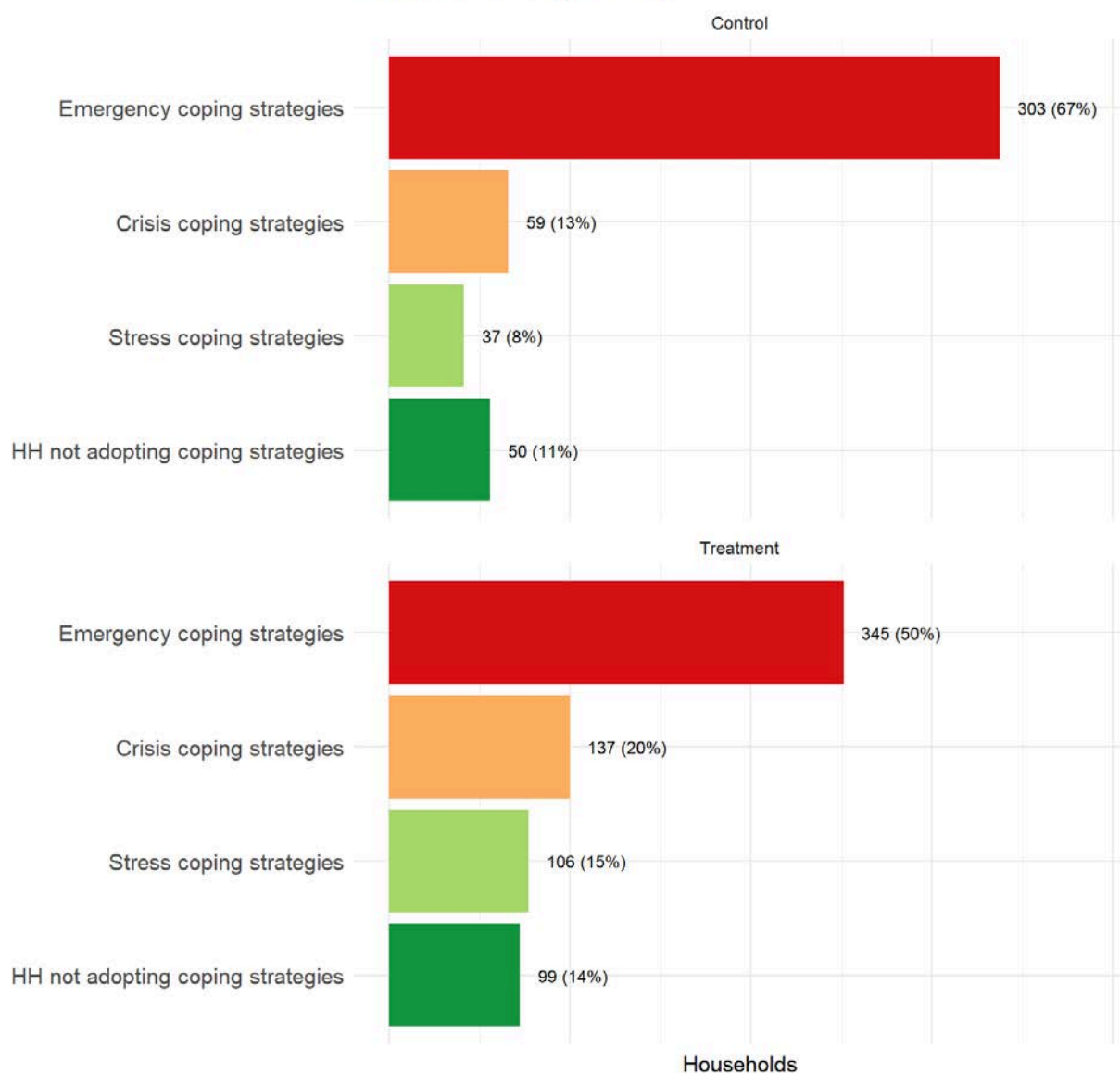
The Livelihood Coping Strategy Index (LCSI) is an important metric used to understand how households cope with food insecurity and livelihood challenges. This data was collected in the latest assessment (December 2023) and categorises coping strategies into four levels based on their severity and potential long-term impacts on household welfare:

1. **HH not adopting coping strategies:** Households that do not resort to any coping strategies, indicating a better food security situation.
2. **Stress coping strategies:** These strategies are less severe and include actions that may affect household well-being but are typically reversible, such as borrowing money or spending savings.
3. **Crisis coping strategies:** More severe than stress strategies, these actions can compromise future productivity, such as selling productive assets or reducing essential expenditures.
4. **Emergency coping strategies:** The most severe strategies, indicating extreme hardship, such as selling the last remaining livestock or engaging in illegal activities.

The plot below presents the distribution of different livelihood coping strategies between control and treatment groups. It visualises the proportions of households adopting each level of coping strategy.

Proportions of Different Livelihood Coping Strategies

Households = 1136 (Dec 2023)



Key observations:

1. Emergency Coping Strategies:
 - **Control Group:** 67% of households in the control group resort to emergency coping strategies.
 - **Treatment Group:** A lower proportion, 50%, of households in the treatment group adopt these strategies.
2. Crisis Coping Strategies:
 - **Control Group:** 13% of households use crisis coping strategies.
 - **Treatment Group:** 20% of households in the treatment group adopt crisis coping strategies.
3. Stress Coping Strategies:
 - **Control Group:** 8% of households employ stress coping strategies.
 - **Treatment Group:** 15% of households in the treatment group resort to these strategies.

4. HH not adopting coping strategies:

- **Control Group:** 11% of households do not adopt any coping strategies.
- **Treatment Group:** 14% of households in the treatment group do not adopt coping strategies.

The higher proportion of households in the treatment group adopting crisis and stress coping strategies compared to the control group suggests a shift from more severe to less severe coping mechanisms. This could indicate some level of improvement in their overall food security and livelihood conditions due to the project intervention.

However, a sizable portion of households still rely on emergency coping strategies, even in the treatment group, indicating ongoing challenges. Continued support and tailored interventions focusing on sustainable livelihood improvements are essential to reduce the need for severe coping strategies further.

Potential recommendations include:

- **Enhanced Livelihood Interventions:** Implement actions that focus on sustainable income-generating activities to reduce dependency on severe coping strategies.
- **Financial Services:** Expand access to cash support, microfinance, and savings groups to help households manage financial shocks without resorting to severe measures.
- **Monitoring and Evaluation:** Continuously monitor the impact of interventions and adjust strategies based on feedback and changing conditions to ensure effectiveness.

These results underscore the importance of sustained and adaptive interventions to improve household resilience and reduce the reliance on detrimental coping strategies.

4.2.3.1 Machine Learning for Identification of LCSi Covariates

To identify the key predictors influencing the Livelihood Coping Strategy Index, we employed a Random Forest model. This approach aids in the selection of covariates for the subsequent Difference-in-Differences analysis, ensuring a robust understanding of the factors impacting LCSi.

The output below shows the performance of the random forest model in predicting the LCSi. Since the LCSi variable `max_coping_behaviourfs` is a categorical variable, the random forest model works as a type of machine learning algorithm used for classification. In this case, it helps us understand how households adopt different coping strategies in response to food insecurity and other challenges. The model uses 500 decision trees to make predictions, and at each split in the decision trees, 28 different variables are considered.

The “OOB (Out-of-Bag) Estimate of Error Rate” is 30.11%. This means that about 30% of the time, the model’s predictions are incorrect when tested on unseen data. While this error rate indicates that the model is not perfect, it still provides valuable insights.

The confusion matrix is a table that shows how well the model classifies the different coping strategies. It compares the actual coping strategies used by households (rows) with the strategies predicted by the model (columns).

The model is particularly good at predicting “Emergency coping strategies,” correctly classifying 610 instances, which is why this category has the lowest error rate. The model has more

difficulty distinguishing between “HH not adopting coping strategies,” “Stress coping strategies,” and “Crisis coping strategies,” as these categories have higher rates of misclassification.

```
Call:
  randomForest(formula = as.factor(max_coping_behaviourfs) ~ .,      data =
df_analysis %>% select(-starts_with("lcsi."), -starts_with("csi."),
-ends_with("_fs")), importance = TRUE)
      Type of random forest: classification
      Number of trees: 500
No. of variables tried at each split: 28

      OOB estimate of  error rate: 30.11%
Confusion matrix:
              HH not adopting coping strategies
HH not adopting coping strategies      66
Stress coping strategies              20
Crisis coping strategies              10
Emergency coping strategies           10

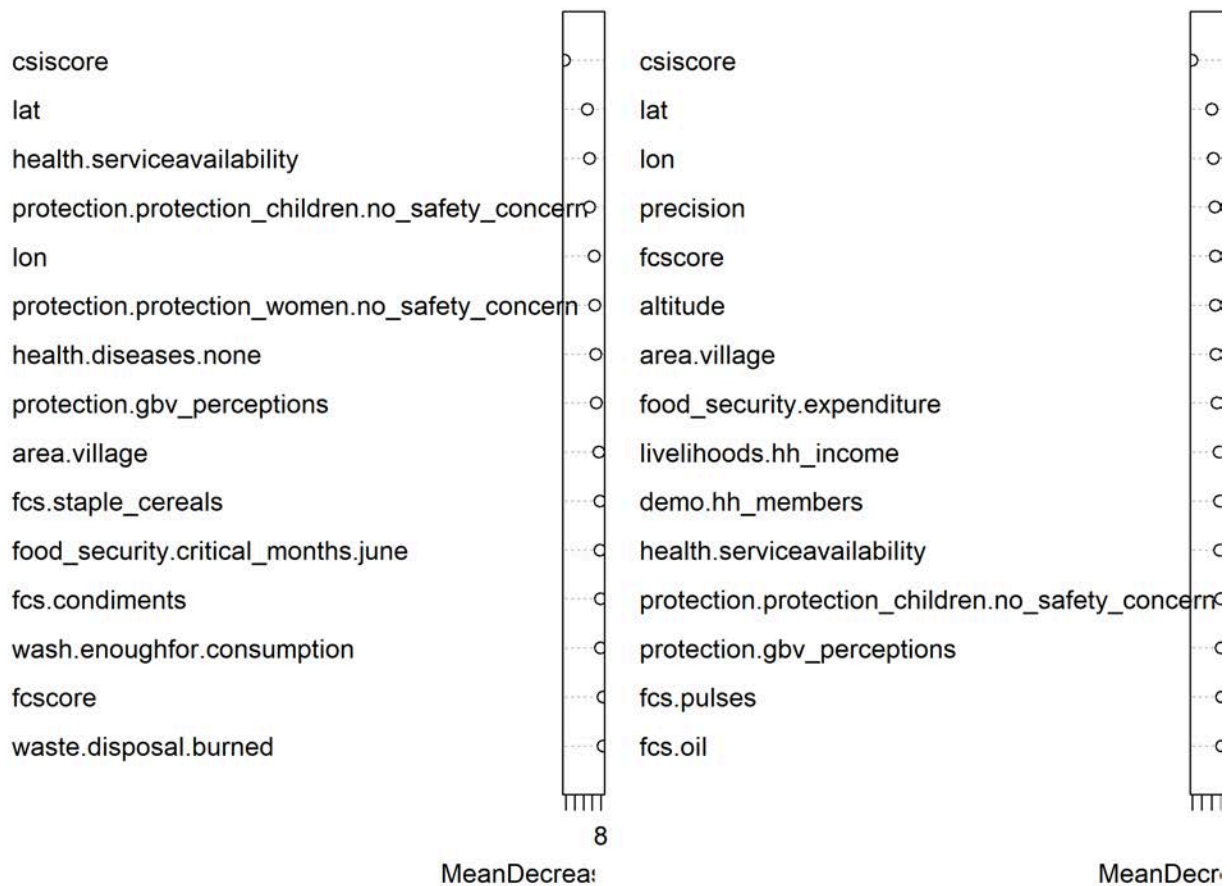
              Stress coping strategies
HH not adopting coping strategies      11
Stress coping strategies              44
Crisis coping strategies               8
Emergency coping strategies           15

              Crisis coping strategies
HH not adopting coping strategies       9
Stress coping strategies                5
Crisis coping strategies              74
Emergency coping strategies            13

              Emergency coping strategies class.error
HH not adopting coping strategies      63  0.55704698
Stress coping strategies              74  0.69230769
Crisis coping strategies             104  0.62244898
Emergency coping strategies           610  0.05864198
```

The graph below illustrates the top 15 predictors of LCSi as determined by the Random Forest model, using the metrics Mean Decrease in Accuracy (%IncMSE) and Mean Decrease in Node Purity (IncNodePurity). These metrics provide insight into the relative importance of numerous factors influencing the LCSi score.

Top 15 predictors of LCSi (random forest model)



Key predictors include:

1. **Coping Strategy Index (csiscore):** The most significant predictor, indicating that households with higher coping strategy indices have higher LCSi scores.
2. **Latitude (lat) and Longitude (lon):** Geographic factors significantly influence LCSi, reflecting regional differences in food security and coping mechanisms.
3. **Health Service Availability (health.serviceavailability):** Availability and quality of health services are crucial in determining how households cope with food insecurity.
4. **Protection Issues:** Variables such as protection.protection_children.no_safety_concern and protection.protection_women.no_safety_concern highlight the importance of safety perceptions in coping strategies.
5. **Food Consumption Scores (fcs):** Different components of food consumption, such as fcs.staple_cereals and fcs.condiments, are important predictors, indicating the relationship between dietary diversity and coping strategies.
6. **Waste Disposal:** Practices such as waste.disposal.burned also influence LCSi, reflecting broader household management strategies.

4.2.3.2 Difference-in-Differences Analysis for LCSi

The DiD model was employed to assess the impact of the SAHARA project on LCSi, incorporating the top five key predictors identified by the Random Forest model. These predictors were included to control for their influence and isolate the effect of the intervention.

```
OLS estimation, Dep. Var.: max_coping_behaviourfs
Observations: 1,136
Standard-errors: Clustered (type)

              Estimate Std. Error   t value
(Intercept)    12.683745    4.811340   2.636219
implementationYear 1 to 3    -0.124959    0.026495  -4.716313
implementationYear 2 to 3 alone    -0.099882    0.025647  -3.894486
csiscore         0.016366    0.003742   4.373325
lat             -0.505649    0.306385  -1.650373
health.serviceavailability.L    -0.167393    0.069066  -2.423662
health.serviceavailability.Q    -0.158224    0.045705  -3.461820
health.serviceavailability.C    -0.451539    0.172717  -2.614322
health.serviceavailability^4    -0.025069    0.083698  -0.299514
health.serviceavailability^5    -0.110572    0.030826  -3.586918
protection.protection_children.no_safety_concern    -0.426928    0.164049  -2.602431
lon             -0.299000    0.111496  -2.681721
Pr(>|t|)
(Intercept)         0.23081
implementationYear 1 to 3         0.13301
implementationYear 2 to 3 alone         0.16001
csiscore             0.14311
lat                 0.34681
health.serviceavailability.L         0.24912
health.serviceavailability.Q         0.17902
health.serviceavailability.C         0.23258
health.serviceavailability^4         0.81474
health.serviceavailability^5         0.17309
protection.protection_children.no_safety_concern    0.23355
lon                 0.22722
... 3 variables were removed because of collinearity (post, implementationYear 1 to
3:post and implementationYear 2 to 3 alone:post)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.945918   Adj. R2: 0.239302
```

The interpretation and key insights from the output above are as follows:

- **Implementation Periods:** The coefficients for `implementationYear 1 to 3` and `implementationYear 2 to 3 alone` are negative, suggesting a reduction in the LCSi score for households within these groups. However, these results are not statistically significant.
- **Coping Strategy Index (csiscore):** This variable has a positive impact on the LCSi score, indicating that households with higher coping strategy indices tend to have higher LCSi scores, reflecting greater food insecurity.
- **Geographic Factors (lat and lon):** These factors significantly influence LCSi, indicating that location plays a crucial role in determining household coping strategies.
- **Health Service Availability:** This variable has a significant negative impact on LCSi, suggesting that better health service availability reduces the reliance on severe coping strategies.

- **Protection Issues:** Safety concerns, particularly those related to children and women, significantly impact LCSi, emphasising the role of protection in household coping mechanisms.

Potential recommendations include:

1. **Targeted Interventions:** Focus on households with high coping strategy indices and significant protection issues, as these are indicators of severe food insecurity and vulnerability.
2. **Health Service Improvements:** Enhance health service availability to reduce the need for severe coping strategies, as indicated by the significant impact of health services on LCSi.
3. **Geographic Tailoring:** Develop region-specific interventions that address the unique challenges faced by households in different geographic locations.
4. **Continuous Monitoring:** Regularly monitor the key predictors identified to assess the effectiveness of interventions and make necessary adjustments.

4.2.3.3 Propensity Score Matching and Regression Analysis for LCSi

We performed a Propensity Score Matching (PSM) analysis to estimate the impact of a project on the Livelihood Coping Strategies Index using data from the fourth assessment. The objective was to balance the treated and control groups based on various covariates and then estimate the treatment effect using a linear regression model.

1. **Data Preparation:** We filtered the dataset to include only the fourth assessment data and selected relevant covariates for the PSM model.
2. **Propensity Score Calculation:** We calculated propensity scores using logistic regression, considering the treatment assignment as the dependent variable and various covariates identified from the random forest importance plot.
3. **Matching:** We applied nearest neighbour matching to pair treated and control units based on their propensity scores.
4. **Balance Check:** We assessed the balance of covariates before and after matching using standardised mean differences, variance ratios, and empirical cumulative distribution function (eCDF) differences.
5. **Linear Regression:** We performed a linear regression on the matched data to estimate the treatment effect on the LCSi, including additional covariates to control for residual confounding.

Call:

```
MatchIt::matchit(formula = psm_formula_lcsi, data = df_analysis,
  method = "nearest", distance = "logit")
```

Summary of Balance for All Data:

	Means Treated	Means Control
distance	0.7530	0.3780
fcscore	31.7336	29.4165
wash.toilet_handwashYes	0.4454	0.3898
wash.toilet_handwashNo	0.4338	0.4855
wash.toilet_handwashNot_applicable	0.1208	0.1247
livelihoods.hh_income	6790.1237	5379.0646
csiscore	34.1485	38.9421

lat	11.9040	11.5661
health.serviceavailabilityVery_good	0.0262	0.0067
health.serviceavailabilityGood	0.2504	0.2116
health.serviceavailabilityRegular	0.2824	0.5167
health.serviceavailabilityBad	0.1965	0.1492
health.serviceavailabilityVery_bad	0.1223	0.0601
health.serviceavailabilityNot_applicable	0.1223	0.0557
protection.protection_children.no_safety_concern	0.6099	0.4120
lon	12.5850	13.0694

Std. Mean Diff. Var. Ratio

distance	3.4892	0.0923
fcscore	0.1818	0.7456
wash.toilet_handwashYes	0.1120	.
wash.toilet_handwashNo	-0.1044	.
wash.toilet_handwashNot_applicable	-0.0120	.
livelihoods.hh_income	0.0783	7.4698
csiscore	-0.2263	0.7562
lat	2.4808	0.1702
health.serviceavailabilityVery_good	0.1222	.
health.serviceavailabilityGood	0.0895	.
health.serviceavailabilityRegular	-0.5205	.
health.serviceavailabilityBad	0.1190	.
health.serviceavailabilityVery_bad	0.1897	.
health.serviceavailabilityNot_applicable	0.2033	.
protection.protection_children.no_safety_concern	0.4057	.
lon	-1.5301	0.2145

eCDF Mean eCDF Max

distance	0.3114	0.4915
fcscore	0.0408	0.1210
wash.toilet_handwashYes	0.0557	0.0557
wash.toilet_handwashNo	0.0518	0.0518
wash.toilet_handwashNot_applicable	0.0039	0.0039
livelihoods.hh_income	0.0359	0.0941
csiscore	0.0455	0.1066
lat	0.2665	0.4566
health.serviceavailabilityVery_good	0.0195	0.0195
health.serviceavailabilityGood	0.0388	0.0388
health.serviceavailabilityRegular	0.2343	0.2343
health.serviceavailabilityBad	0.0473	0.0473
health.serviceavailabilityVery_bad	0.0621	0.0621
health.serviceavailabilityNot_applicable	0.0666	0.0666
protection.protection_children.no_safety_concern	0.1979	0.1979
lon	0.1772	0.4566

Summary of Balance for Matched Data:

Means Treated Means Control

distance	0.8176	0.3780
fcscore	32.4788	29.4165
wash.toilet_handwashYes	0.4410	0.3898
wash.toilet_handwashNo	0.4098	0.4855
wash.toilet_handwashNot_applicable	0.1492	0.1247
livelihoods.hh_income	7942.3163	5379.0646
csiscore	30.8107	38.9421
lat	11.9619	11.5661
health.serviceavailabilityVery_good	0.0401	0.0067
health.serviceavailabilityGood	0.1871	0.2116
health.serviceavailabilityRegular	0.1915	0.5167
health.serviceavailabilityBad	0.2584	0.1492
health.serviceavailabilityVery_bad	0.1782	0.0601
health.serviceavailabilityNot_applicable	0.1448	0.0557
protection.protection_children.no_safety_concern	0.5657	0.4120

lon	12.7092	13.0694
	Std. Mean	Diff. Var. Ratio
distance	4.0911	0.0289
fcscore	0.2403	0.7734
wash.toilet_handwashYes	0.1031	.
wash.toilet_handwashNo	-0.1528	.
wash.toilet_handwashNot_applicable	0.0752	.
livelihoods.hh_income	0.1423	11.2110
csiscore	-0.3839	0.7499
lat	2.9057	0.1362
health.serviceavailabilityVery_good	0.2091	.
health.serviceavailabilityGood	-0.0566	.
health.serviceavailabilityRegular	-0.7223	.
health.serviceavailabilityBad	0.2746	.
health.serviceavailabilityVery_bad	0.3603	.
health.serviceavailabilityNot_applicable	0.2719	.
protection.protection_children.no_safety_concern	0.3151	.
lon	-1.1378	0.1735
	eCDF Mean	eCDF Max
distance	0.4525	0.7661
fcscore	0.0475	0.1359
wash.toilet_handwashYes	0.0512	0.0512
wash.toilet_handwashNo	0.0757	0.0757
wash.toilet_handwashNot_applicable	0.0245	0.0245
livelihoods.hh_income	0.0402	0.0735
csiscore	0.0782	0.1670
lat	0.3674	0.5880
health.serviceavailabilityVery_good	0.0334	0.0334
health.serviceavailabilityGood	0.0245	0.0245
health.serviceavailabilityRegular	0.3252	0.3252
health.serviceavailabilityBad	0.1091	0.1091
health.serviceavailabilityVery_bad	0.1180	0.1180
health.serviceavailabilityNot_applicable	0.0891	0.0891
protection.protection_children.no_safety_concern	0.1537	0.1537
lon	0.1471	0.4566
	Std. Pair	Dist.
distance	4.0911	
fcscore	1.2483	
wash.toilet_handwashYes	0.9903	
wash.toilet_handwashNo	0.9977	
wash.toilet_handwashNot_applicable	0.7859	
livelihoods.hh_income	0.4508	
csiscore	1.2438	
lat	3.0505	
health.serviceavailabilityVery_good	0.2928	
health.serviceavailabilityGood	0.7660	
health.serviceavailabilityRegular	1.1379	
health.serviceavailabilityBad	0.7791	
health.serviceavailabilityVery_bad	0.6459	
health.serviceavailabilityNot_applicable	0.5711	
protection.protection_children.no_safety_concern	1.0821	
lon	2.3741	

Sample Sizes:

	Control	Treated
All	449	687
Matched	449	449
Unmatched	0	238
Discarded	0	0

The output summary can be summarised as follows:

Balance for All Data: - The standardised mean differences (Std. Mean Diff.) and variance ratios (Var. Ratio) for covariates such as distance, fcscore, wash.toilet_handwashYes, livelihoods.hh_income, csiscore, and lat indicate varying degrees of imbalance before matching. - High Std. Mean Diff. and Var. Ratio values, especially for distance and lat, suggest significant initial imbalance.

Balance for Matched Data: - After matching, most covariates show improved balance, but some, such as distance and lat, still exhibit substantial differences. - eCDF Mean and eCDF Max values also indicate improved, though not perfect, balance across covariates. - Sample Sizes: Matched data include 449 treated and 449 control units, with 238 treated units unmatched due to lack of suitable control matches.

The PSM analysis shows that matching significantly improves the balance between treated and control groups, although some covariates still exhibit notable differences. The matched dataset was used for further regression analysis to estimate the project's impact on LCSI, ensuring a more reliable comparison of treated and control groups by mitigating initial imbalances.

By balancing the covariates, the PSM allows for a more accurate estimation of the treatment effect, helping to understand how the project impacts household coping strategies more clearly.

Call:

```
lm(formula = as.formula(paste("max_coping_behaviourfs ~ treatment + fcscore + food_security.expenditure + wash.toilet_handwash + \nwash.watertreat_needed + wash.toilet_separate + wash.shortage + \nwaste.disposalchange + protection.safety_felling + livelihoods.hh_income + ", paste(top_vars_lcsi, collapse = " + "))), data = matched_data)
```

Residuals:

LABEL: Summary of asset depletion

VALUES:

-2.83861, -0.58744, 0.13964, 0.68655, 2.30536

VALUE LABELS:

- 1 HH not adopting coping strategies
- 2 Stress coping strategies
- 3 Crisis coping strategies
- 4 Emergency coping strategies

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.678e+01	3.042e+00	5.516
treatment	-7.548e-02	7.490e-02	-1.008
fcscore	-9.176e-03	2.589e-03	-3.544
food_security.expenditure	7.693e-07	1.516e-06	0.508
wash.toilet_handwash.L	-8.983e-02	8.934e-02	-1.006
wash.toilet_handwash.Q	-1.386e-01	6.520e-02	-2.125
wash.watertreat_needed.L	-1.422e-01	9.114e-02	-1.560
wash.watertreat_needed.Q	-2.874e-01	6.697e-02	-4.292
wash.toilet_separate.L	-1.861e-01	8.761e-02	-2.125
wash.toilet_separate.Q	3.930e-02	6.078e-02	0.647
wash.shortage.L	-1.463e-01	1.607e-01	-0.910
wash.shortage.Q	1.039e-01	1.004e-01	1.035
waste.disposalchange.L	8.065e-02	1.045e-01	0.772
waste.disposalchange.Q	1.961e-01	6.735e-02	2.911
protection.safety_felling.L	-2.647e-01	3.634e-01	-0.728

protection.safety_felling.Q	-1.404e-01	3.022e-01	-0.464
protection.safety_felling.C	-2.646e-02	2.352e-01	-0.113
protection.safety_felling^4	2.969e-02	1.437e-01	0.207
livelihoods.hh_income	1.972e-06	1.961e-06	1.006
csiscore	1.825e-02	1.530e-03	11.923
lat	-8.778e-01	2.064e-01	-4.254
health.serviceavailability.L	-2.222e-02	1.631e-01	-0.136
health.serviceavailability.Q	-5.529e-02	1.385e-01	-0.399
health.serviceavailability.C	-3.661e-01	1.145e-01	-3.196
health.serviceavailability^4	6.083e-03	9.268e-02	0.066
health.serviceavailability^5	-1.495e-01	7.211e-02	-2.073
protection.protection_children.no_safety_concern	-2.240e-01	7.243e-02	-3.093
lon	-2.868e-01	9.597e-02	-2.989
	Pr(> t)		
(Intercept)	4.58e-08	***	
treatment	0.313823		
fcscore	0.000415	***	
food_security.expenditure	0.611864		
wash.toilet_handwash.L	0.314932		
wash.toilet_handwash.Q	0.033843	*	
wash.watertreat_needed.L	0.119068		
wash.watertreat_needed.Q	1.97e-05	***	
wash.toilet_separate.L	0.033909	*	
wash.toilet_separate.Q	0.517998		
wash.shortage.L	0.362948		
wash.shortage.Q	0.300825		
waste.disposalchange.L	0.440539		
waste.disposalchange.Q	0.003696	**	
protection.safety_felling.L	0.466524		
protection.safety_felling.Q	0.642452		
protection.safety_felling.C	0.910443		
protection.safety_felling^4	0.836352		
livelihoods.hh_income	0.314921		
csiscore	< 2e-16	***	
lat	2.33e-05	***	
health.serviceavailability.L	0.891677		
health.serviceavailability.Q	0.689942		
health.serviceavailability.C	0.001442	**	
health.serviceavailability^4	0.947682		
health.serviceavailability^5	0.038427	*	
protection.protection_children.no_safety_concern	0.002045	**	
lon	0.002879	**	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9198 on 870 degrees of freedom

Multiple R-squared: 0.306, Adjusted R-squared: 0.2844

F-statistic: 14.21 on 27 and 870 DF, p-value: < 2.2e-16

The summary of the regression analysis above is as follows:

- **Treatment:** The treatment effect was not statistically significant (p-value = 0.314), suggesting no robust evidence of a treatment effect on the LCSl.
- **CSI Score (csiscore):** Positive and significant impact on LCSl, indicating that higher coping strategy index scores are associated with higher LCSl (crisis and emergency).
- **Food Consumption Score (fcscore):** Negative and significant impact on LCSl, suggesting that higher food consumption scores are associated with lower LCSl, as expected.

- **WASH Variables:**
 - wash.watertreat_needed: Negative and significant, indicating households needing water treatment tend to have lower LCSi.
 - wash.toilet_separate: Positive and significant, suggesting that households with separate toilets tend to have higher LCSi.
- **Protection Variables:**
 - protection.protection_children.no_safety_concern: Negative and significant, suggesting that households perceiving no safety concerns for children tend to have lower LCSi.
- **Geographical Factors:**
 - lat: Negative and significant, indicating that geographical location (latitude) influences LCSi.
- **Health Service Availability:** Mixed results, with some categories showing significant negative impacts on LCSi.

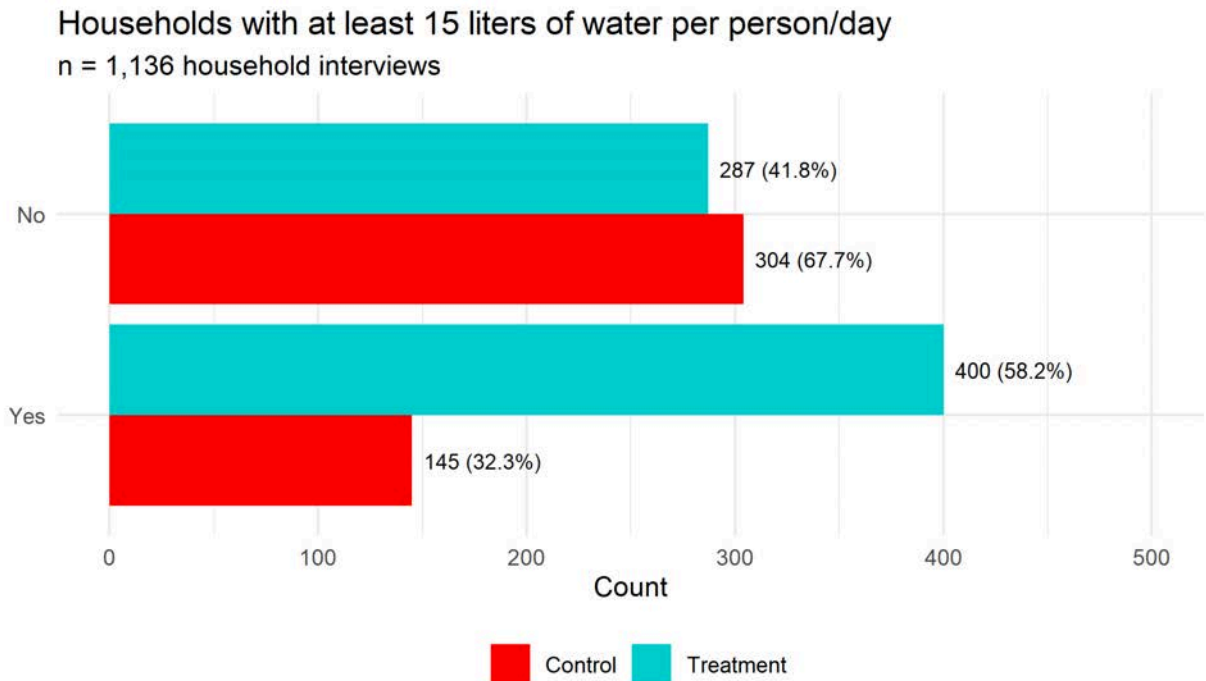
The PSM and subsequent linear regression analysis provided insights into the factors affecting the LCSi. Despite the improved balance after matching, the treatment effect was not statistically significant, suggesting no robust evidence of the project's impact on LCSi. However, other factors such as coping strategy index scores, food consumption scores, and various WASH-related variables showed significant associations with the LCSi. This analysis underscores the importance of considering multiple factors when evaluating the impact of development projects on household coping strategies.

Recommendations:

1. **Focus on WASH Interventions:**
 - Ensure access to clean water and promote water treatment practices.
 - Intensify efforts to improve sanitation infrastructure by providing separate toilet facilities.
 - Address water shortages through reliable water supply systems and conservation practices.
 - Promote hygiene education and provision of hygiene products.
2. **Enhance Food Security Measures:**
 - Support agricultural productivity, food distribution systems, and nutritional programs to improve food consumption scores.
3. **Support for Household Income Generation:**
 - Implement vocational training, microfinance, and employment programs to improve household income and resilience.
4. **Integrated Waste Management:**
 - Promote sustainable waste management systems, including recycling and proper disposal facilities.
5. **Community and Social Protection Programmes:**
 - Enhance community security and protection, especially for IDPs and vulnerable populations, through community policing and social protection programs.

4.2.4 Access to water

This section explores impacts in terms of access to water, considering the variable `wash.percapita` which refers to answers to the question *Is your household accessing at least 15 litres of water per person per day?* Respondents could answer “Yes”, “No” and “Not applicable” to this question. For simplicity, we recoded those answers “Not applicable” to “No”, as they make 0.8 out of the 1,136 total observations (0.8%).



The bar plot shows the distribution of households that access at least 15 litres of water per person per day, separated by the “Treatment” and “Control” groups. Key takeaways are:

1. **Access to Water:** A higher percentage of households in the “Treatment” group (58%) reported accessing at least 15 litres of water per person per day compared to the “Control” group (32%).
2. **Water Shortage:** Conversely, a higher percentage of households in the “Control” group (68%) reported not having access to the minimum 15 litres of water per person per day compared to the “Treatment” group (42%).

This suggests that interventions in the “Treatment” group may have been more effective in ensuring adequate water access per capita. This difference highlights the potential positive impact of the implemented water-related measures in improving water access for households in the “Treatment” group.

4.2.4.1 Machine learning to identify key covariates for water access

We employed a Random Forest model to identify the key covariates that predict whether households are accessing at least 15 litres of water per person per day. The dataset includes a range of factors related to household demographics, livelihoods, WASH (Water, Sanitation, and Hygiene) conditions, and other variables. Below, we present and explain the results of the Random Forest model based on the variable importance measures.

As shown in the output below, the Random Forest model was trained using 500 trees. The model's Mean of Squared Residuals was calculated, indicating the model's performance in predicting the outcome variable.

```
Call:
randomForest(formula = wash.percapita ~ ., data = df_analysis, importance = TRUE)
```

Type of random forest: classification

Number of trees: 500

No. of variables tried at each split: 29

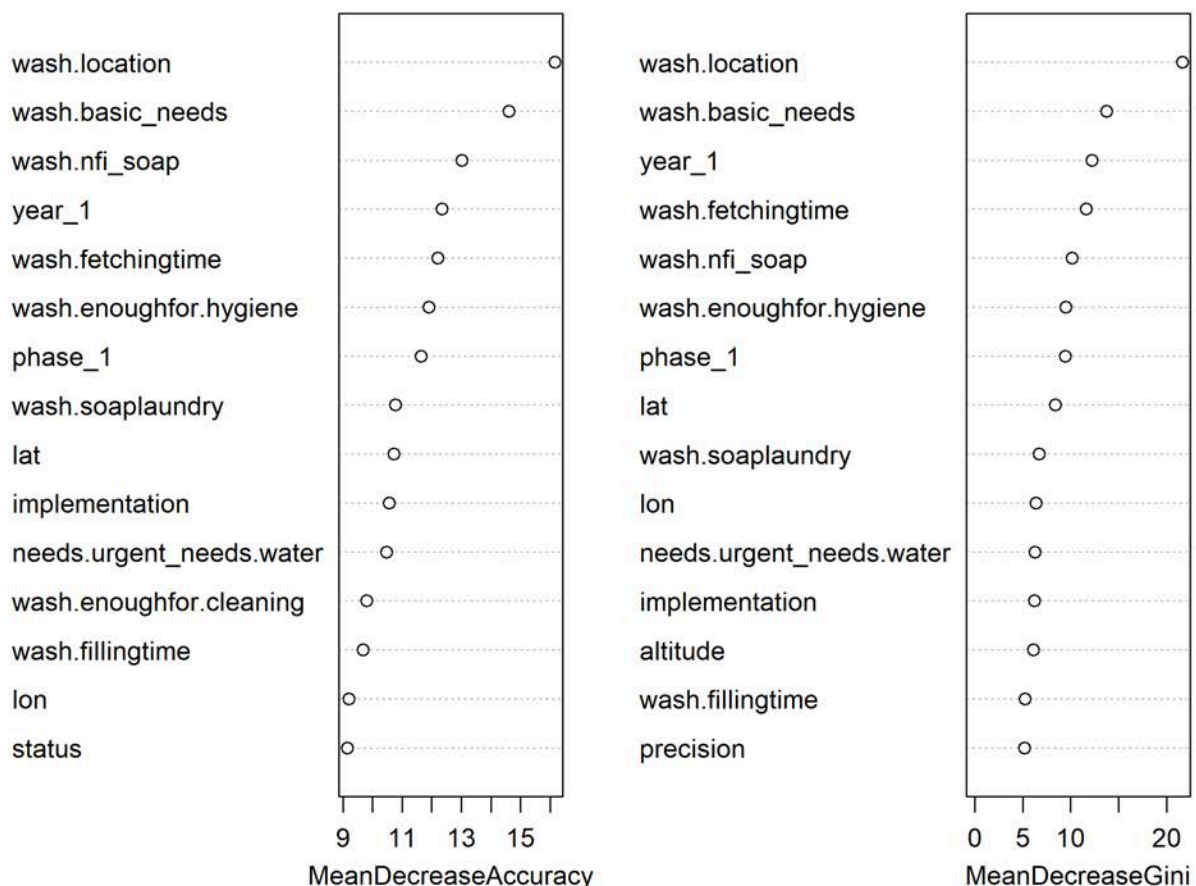
OOB estimate of error rate: 12.15%

Confusion matrix:

	Yes	No	class.error
Yes	467	78	0.1431193
No	60	531	0.1015228

The importance of each variable (see plot below) was assessed using two metrics: the Mean Decrease in Mean Squared Error (%IncMSE) and the Mean Decrease in Node Purity (IncNodePurity). These metrics provide insights into which variables are the most influential in predicting the response variable, which in this case is wash.percapita.

Top 15 predictors of water access (random forest model)



The importance of each variable in predicting water access was assessed using two metrics: Mean Decrease in Accuracy and Mean Decrease in Gini (Node Purity). These metrics highlight the most influential factors affecting water access.

The main results can be summarised as follows:

1. **wash.location:** The location of the water source is the most critical factor in determining water access per capita. Households closer to water sources have better access.
2. **wash.basic_needs:** The ability to meet basic water needs significantly impacts water access, indicating households' overall capability to secure necessary water resources.
3. **wash.fetchingtime:** The time taken to fetch water is crucial, with shorter fetching times typically correlating with better access to water.
4. **wash.nfi_soap:** The availability of non-food items such as soap is essential, highlighting the importance of hygiene resources in ensuring water access.
5. **year_1:** The first year of implementation shows a significant impact, indicating changes over time and the importance of project stages.
6. **lat and lon:** Geographic coordinates (latitude and longitude) significantly affect water access, reflecting the importance of the physical location of households.
7. **implementation:** The phase of project implementation impacts water access, showing that ongoing or recently completed interventions are influential.

The analysis identifies the location of water sources, basic water needs, fetching time, availability of hygiene items, implementation phases, and geographic location as the most critical factors influencing water access.

Policy recommendations based on random forest model results

1. **Improve Proximity to Water Sources:** Infrastructure projects should focus on reducing the distance between households and water sources. Bringing water sources closer to residential areas can significantly enhance water access.
2. **Enhance Distribution of Hygiene Supplies:** Ensuring the availability of hygiene-related non-food items such as soap is crucial. Hygiene promotion programmes should be prioritised to improve overall water access.
3. **Address Basic Water Needs:** Policies should aim to secure the basic water needs of households, ensuring sufficient water for daily use.

4.2.4.2 PSM and regression analysis for water access

After the PSM matching process to refine the comparability of treatment and control households, we used logistic regression to explore the impacts and key predictors of water access. The output below summarises the results of that regression model.

Call:

```
lm(formula = psm_formula_water, data = matched_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.01813	-0.23662	0.01084	0.22119	1.03624

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.95412	0.04781	40.873	< 2e-16 ***
treatment	-0.10511	0.02965	-3.545	0.000413 ***
wash.location.L	0.67768	0.09132	7.421	2.73e-13 ***
wash.location.Q	0.18889	0.05506	3.431	0.000630 ***
wash.basic_needs.L	0.33221	0.06824	4.868	1.34e-06 ***
wash.basic_needs.Q	-0.03436	0.05762	-0.596	0.551078 .
wash.basic_needs.C	-0.08148	0.04671	-1.744	0.081436 .
wash.basic_needs^4	0.01460	0.03581	0.408	0.683566 .
wash.basic_needs^5	-0.03603	0.02741	-1.314	0.189086 .
wash.nfi_soap.L	0.16114	0.08327	1.935	0.053288 .
wash.nfi_soap.Q	-0.10778	0.04884	-2.207	0.027598 *
year_1Yes	-0.22970	0.03444	-6.670	4.50e-11 ***
wash.fetchingtime.L	0.09794	0.03138	3.121	0.001859 **
wash.fetchingtime.Q	-0.03683	0.02571	-1.432	0.152401 .
wash.fetchingtime.C	-0.04189	0.02491	-1.681	0.093042 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3683 on 883 degrees of freedom
Multiple R-squared: 0.493, Adjusted R-squared: 0.4849
F-statistic: 61.32 on 14 and 883 DF, p-value: < 2.2e-16

The output above provides elements for analysing water access based on the regression analysis with PSM-matched data.

Intervention Effect: The coefficient for the treatment variable is statistically significant (p-value = 0.000413), indicating that the project intervention had a significant direct impact on households' access to water.

The variable wash.percapita represents whether households have access to at least 15 litres of water per person per day. It is coded as follows: - **Yes (1)**: Household has access to the required water per capita. - **No (2)**: Household does not have access to the required water per capita. - **Not_applicable (3)**: This category is less relevant in this context but included for completeness.

The project intervention's coefficient is negative and statistically significant, indicating that, on average, households that received the treatment are slightly less likely to have access to at least 15 litres of water per person per day compared to those that did not receive the treatment. Reasons include prioritising the most vulnerable areas, for example. Initial implementation challenges may also have an impact. During the initial stages of the project, there might have been logistical or operational challenges that temporarily hindered water access improvements. In some cases, the intervention might have focused on other urgent needs or longer-term infrastructure projects that did not immediately improve water access but could have other long-term benefits. Measurement timing may also play a role since the impact assessment might have occurred during a period of transition or adjustment, where the full benefits of the intervention had not yet been realised.

Important Covariates:

1. Water Source Location (wash.location):

- The location of the water source is critical. Households with better proximity to water sources have significantly improved access (p < 0.001).

2. Basic Water Needs (wash.basic_needs):
 - Households that meet their basic water needs are more likely to have sufficient water access. This factor shows a significant positive association ($p < 0.001$).
3. Fetching Time (wash.fetchingtime):
 - Shorter water fetching times are associated with better water access, highlighting the importance of reducing the distance or time required to collect water ($p = 0.001859$).
4. Availability of Soap (wash.nfi_soap):
 - The availability of non-food items such as soap also plays a role in water access, with mixed effects. The linear term shows a marginally significant positive association ($p = 0.053288$), while the quadratic term is negatively associated ($p = 0.027598$).
5. Year 1 Implementation (year_1Yes):
 - The implementation phase also significantly affects water access, with a negative coefficient indicating that initial implementation stages might have faced challenges ($p < 0.001$).

These results provide a clear understanding of the factors influencing water access and highlight the importance of targeted interventions to improve water availability and quality for households. The significant impact of the treatment variable underscores the effectiveness of the project intervention in enhancing water access.

Potential Policy Recommendations

1. **Improve Proximity to Water Sources:** Reduce the distance between households and water sources through infrastructure projects that bring water sources closer to residential areas.
2. **Enhance Distribution of Hygiene Supplies:** Ensure the availability of non-food items such as soap to improve overall water access. Prioritise hygiene promotion programs.
3. **Address Basic Water Needs:** Secure the basic water needs of households by providing sufficient water for daily use.
4. **Reduce Water Fetching Times:** Implement measures to reduce the time required to collect water, such as closer water points and efficient water distribution systems.
5. **Monitor Implementation Phases:** Continuously monitor and evaluate the impact of different implementation phases on water access to optimise the timing and sequencing of interventions.

These recommendations aim to improve water access by addressing the critical factors identified in the analysis, ensuring that interventions are effective and targeted to the needs of the households.

4.2.4.3 DiD Model for Water Access

A DiD model was employed to assess the impact of the project intervention on households' access to at least 15 litres of water per person per day (wash.percapita). This model includes the interaction term between the implementation period and the post-intervention period, along with various covariates identified as significant in previous analyses.

```

OLS estimation, Dep. Var.: as.numeric(wash.percapita)
Observations: 898
Standard-errors: Clustered (type)

              Estimate Std. Error   t value Pr(>|t|)
(Intercept)      1.954123    0.032079   60.91542 0.010450 *
implementationYear 1 to 3      -0.334804    0.023675  -14.14169 0.044942 *
implementationYear 2 to 3 alone -0.105108    0.002546  -41.28901 0.015416 *
wash.location.L         0.677680    0.064761   10.46428 0.060653 .
wash.location.Q         0.188887    0.065654    2.87701 0.212962
wash.basic_needs.L       0.332209    0.071544    4.64345 0.135038
wash.basic_needs.Q      -0.034364    0.011195   -3.06947 0.200501
wash.basic_needs.C      -0.081479    0.038149   -2.13578 0.278773
wash.basic_needs^4       0.014602    0.007738    1.88712 0.310216
wash.basic_needs^5      -0.036029    0.015697   -2.29527 0.261575
wash.nfi_soap.L         0.161142    0.050489    3.19161 0.193299
wash.nfi_soap.Q        -0.107781    0.004557  -23.65379 0.026898 *
wash.fetchingtime.L      0.097943    0.006568   14.91158 0.042629 *
wash.fetchingtime.Q     -0.036826    0.023549   -1.56379 0.362197
wash.fetchingtime.C     -0.041889    0.002557  -16.37908 0.038820 *
... 4 variables were removed because of collinearity (post, year_1Yes and 2 others
[full set in $collin.var])
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.365171  Adj. R2: 0.484935

```

Key findings from the DiD model are as follows:

1. Intervention Effect:

- The coefficients for the treatment periods (implementationYear 1 to 3 and implementationYear 2 to 3 alone) are negative and statistically significant, suggesting that during these periods, there were challenges in ensuring water access, compared to households in areas not covered by the project. Specifically:
 - implementationYear 1 to 3: Estimate = -0.334804, p-value = 0.044942
 - implementationYear 2 to 3 alone: Estimate = -0.105108, p-value = 0.015416

2. Significant Covariates:

- **Water Source Location** (wash.location.L): A positive and statistically significant association with water access, indicating that closer proximity to water sources improves water availability. (Estimate = 0.677680, p-value = 0.060653)
- **Fetching Time** (wash.fetchingtime.L): A positive association, suggesting shorter fetching times are beneficial for water access. (Estimate = 0.097943, p-value = 0.042629)
- **Availability of Soap** (wash.nfi_soap.Q): A negative association, indicating that households with sufficient soap might face less water access issues. (Estimate = -0.107781, p-value = 0.026898)

Potential Policy Recommendations:

1. Enhance Proximity to Water Sources:

- **Infrastructure Projects:** Prioritise reducing the distance between households and water sources. Building and maintaining nearby water infrastructure can significantly enhance water access.

2. Improve Water Fetching Efficiency:

- **Shorten Fetching Times:** Implement strategies to reduce the time spent fetching water, such as increasing the number of water points and improving transportation means.
- 3. Promote Hygiene Supplies:
 - **Distribution of Soap:** Ensure the continuous availability of hygiene supplies, such as soap, which has shown to have a mixed but positive impact on water access.
- 4. Address Implementation Challenges:
 - **Evaluate and Adjust:** Regularly evaluate the project implementation phases to address and mitigate any challenges that arise, ensuring continuous improvement and adaptation of strategies to improve water access.

4.2.5 Protection

4.2.5.1 Machine learning to identify key covariates for safety feeling

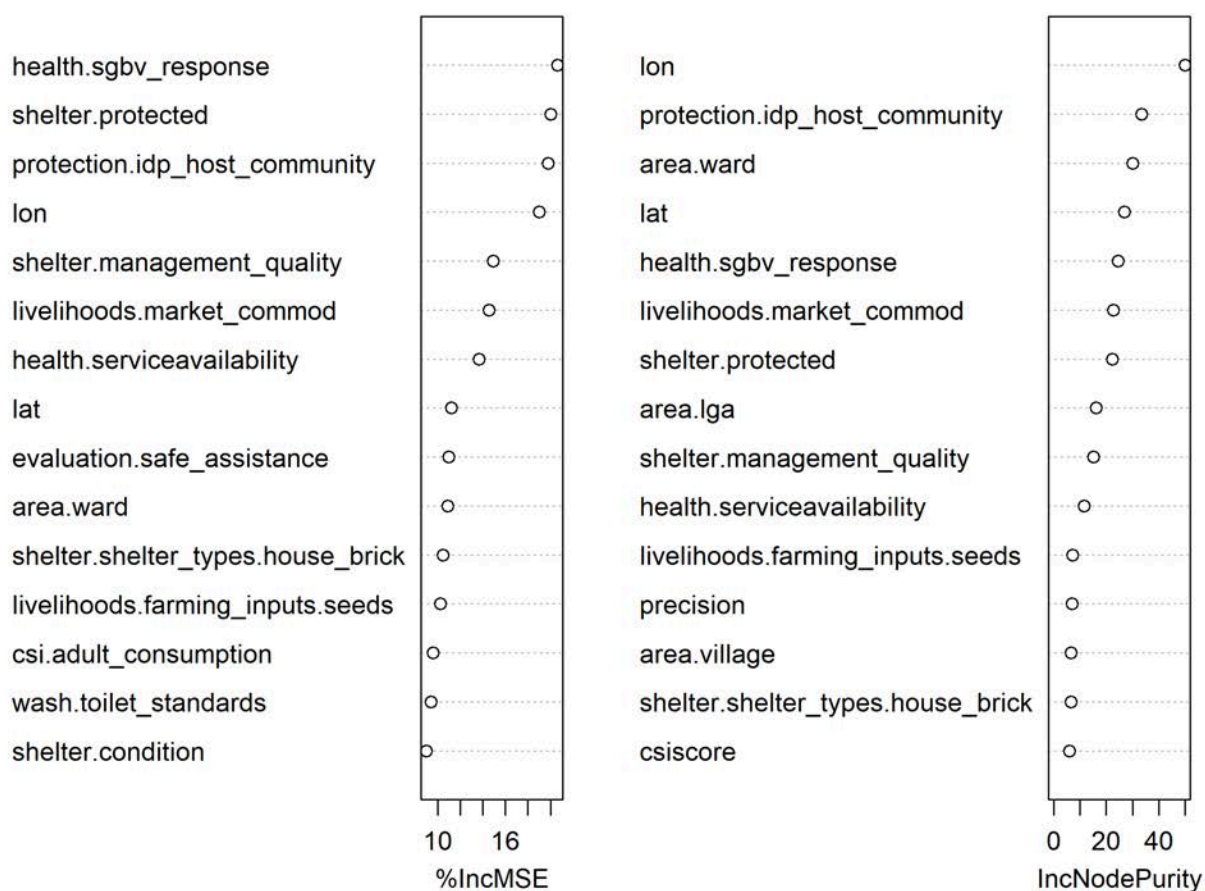
To identify the key covariates influencing the sense of safety among households, we employed a Random Forest regression model. The model aimed to predict the `protection.safety_felling` variable using a comprehensive set of predictors. The summary of the model is as follows:

```
Call:
  randomForest(formula = protection.safety_felling ~ ., data = df_analysis_rf,
    importance = TRUE)
      Type of random forest: regression
      Number of trees: 500
No. of variables tried at each split: 287

      Mean of squared residuals: 0.236482
      % Var explained: 59.99
```

The importance of each variable was assessed using two metrics: Mean Decrease in Mean Squared Error (%IncMSE) and Mean Decrease in Node Purity (IncNodePurity). These metrics provide insights into which variables most significantly influence the sense of safety among households.

Key predictors of safety feeling (protection)



The importance of each variable was assessed using two metrics: Mean Decrease in Mean Squared Error (%IncMSE) and Mean Decrease in Node Purity (IncNodePurity). These metrics provide insights into which variables most significantly influence the sense of safety among households.

Key Predictors of Safety Feeling: - Geographic Location (lon, lat): Geographic coordinates were significant predictors, indicating the importance of location in determining safety feelings.

- **Protection Factors:** Variables related to the quality of protection and community relationships, such as protection.idp_host_community and health.sgbv_response, were critical in influencing safety.
- **Shelter Conditions:** The quality and type of shelter, including management quality and construction type, were significant predictors.
- **Health and Service Availability:** Access to health services and responses to gender-based violence (GBV) significantly impacted feelings of safety.
- **Livelihood and Market Conditions:** Variables related to market commodities and livelihoods also played a role in determining safety perceptions.

4.2.5.2 DiD Model for safety feeling

To further assess the impact of the project intervention on households' sense of safety, we employed a Difference-in-Differences model. This model included the interaction term between the implementation period and the post-intervention period, along with various covariates identified as significant in the Random Forest analysis.

```
OLS estimation, Dep. Var.: as.numeric(protection.safety_felling)
Observations: 898
Standard-errors: Clustered (type)

      Estimate Std. Error   t value
(Intercept)      -16.126997  33.571001  -0.480385
implementationYear 1 to 3      -0.108744   0.007856 -13.841822
implementationYear 2 to 3 alone    -0.134429   0.027810  -4.833920
lon              0.247931   2.155753   0.115009
protection.idp_host_communityGood    -0.518744   0.514286  -1.008668
protection.idp_host_communityNot_applicable -0.361571   0.484279  -0.746618
protection.idp_host_communityRegular  -0.214984   0.533285  -0.403132
protection.idp_host_communityVery_bad  -0.048331   0.572235  -0.084459
protection.idp_host_communityVery_good -0.928435   0.533545  -1.740124
area.ward.L        1.263105  10.410867   0.121326
area.ward.Q         0.138999   8.690213   0.015995
area.ward.C       -0.333336   3.912698  -0.085193
lat               1.356177   0.257424   5.268267
health.sgbv_response.L      0.629867   0.123445   5.102391
health.sgbv_response.Q    -0.661160   0.082201  -8.043233
health.sgbv_response.C    -0.134832   0.094106  -1.432761
health.sgbv_response^4    -0.023747   0.106005  -0.224016
health.sgbv_response^5     0.062426   0.033500   1.863472

Pr(>|t|)
(Intercept)      0.714901
implementationYear 1 to 3      0.045913 *
implementationYear 2 to 3 alone    0.129867
lon              0.927103
protection.idp_host_communityGood    0.497253
protection.idp_host_communityNot_applicable 0.591715
protection.idp_host_communityRegular  0.756045
protection.idp_host_communityVery_bad  0.946359
protection.idp_host_communityVery_good  0.332053
area.ward.L        0.923137
area.ward.Q         0.989818
area.ward.C         0.945895
lat               0.119420
health.sgbv_response.L      0.123207
health.sgbv_response.Q      0.078746 .
health.sgbv_response.C      0.387925
health.sgbv_response^4      0.859703
health.sgbv_response^5      0.313550
... 4 variables were removed because of collinearity (post, area.ward^4 and 2 others
[full set in $collin.var])
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.598039  Adj. R2: 0.300783
```

Key findings from the DiD model include:

1. **Intervention Effect:**

- **Implementation Year 1 to 3:** The coefficient is negative (-0.108744) and statistically significant ($p = 0.045913$), indicating that during the first to third years of implementation, there was an improvement in safety feelings. A lower score on the `protection.safety_felling` variable (where 1 = Very good and 5 = Very bad) corresponds to better safety feelings.
- **Implementation Year 2 to 3 alone:** This coefficient is also negative (-0.134429) but not statistically significant ($p = 0.129867$), suggesting a trend towards improved safety feelings during the second to third years alone, though the effect is not strong enough to be conclusive.

2. **Significant Covariates:**

- **Latitude (`lat`):** Positive and significant association ($t = 5.268267$, $p < 0.001$), suggesting that households located further north report better safety feelings.
- **GBV Response (`health.sgbv_response.L`):** Positive association ($t = 5.102391$, $p < 0.001$) indicating improved safety feelings. This means that better GBV response mechanisms are associated with improved safety perceptions.
- **GBV Response (`health.sgbv_response.Q`):** Negative association ($t = -8.043233$, $p < 0.001$) with quadratic term, highlighting the complexity of the impact of GBV response on safety feelings, highlighting that some aspects of the GBV response might negatively affect safety feelings.

Potential Policy Recommendations:

1. **Consistent and Long-Term Interventions:** Given the significant improvement in safety feelings during the first to third years, continuous and consistent interventions should be prioritised to maintain and enhance safety.
2. **Tailored Interventions for Different Regions:** The positive association with latitude suggests the need for region-specific strategies to improve safety feelings, particularly focusing on northern regions.
3. **Invest in Effective GBV Responses:** Strengthen and improve the mechanisms addressing gender-based violence to boost community safety feelings.
4. **Improve IDP Host Community Relations:** Promote positive interactions and relations between internally displaced persons (IDPs) and host communities to enhance overall safety.
5. **Improve Shelters and Health Access:** Enhance shelter quality and health service availability to contribute to better safety feelings in the community.

5 Conclusion and Recommendations

The independent impact analysis of the SAHARA project in Borno State, Nigeria, has revealed substantial positive changes across several dimensions of community well-being. The evaluation employed Difference-in-Differences models, machine learning techniques, Propensity Score Matching, and various statistical tests to derive the following key findings:

Village-Level Impact:

1. **Vulnerability Score:**
 - The project significantly reduced vulnerability scores, particularly in villages involved from January 2021 to December 2022/January 2023 and those involved only from December 2021/January 2022 to December 2023. This indicates a sustained and increasing positive impact post-intervention.
2. **Key Indicators:**
 - **Food Security (FCS):** There were significant improvements, especially in villages involved from Year 1 to 3, with a notable positive impact on households with unaccompanied children.
 - **Coping Strategies (CSI):** There were improvements in coping strategies, particularly post-intervention, suggesting reduced reliance on negative coping mechanisms.
 - **Safety Feeling and Safe Assistance:** There were significant improvements in safety feelings and perceptions of safe assistance delivery.
 - **Water Access and Sufficiency:** There were significant improvements in water access, proximity to water sources, and water sufficiency.
 - **Body Soap Access:** There were notable improvements in access to body soap, particularly in the Year 1 to 3 group.

Household-Level Impact:

1. **Food Security (FCS):** Significant reductions in households below the acceptable FCS threshold were observed in both implementation groups. Despite initial instability in food consumption, there were sustained improvements in the treatment group by the final assessment.
2. **Coping Strategies (CSI):** The project significantly reduced CSI scores, especially in the long term. Initial increases in CSI scores during early intervention phases were followed by gradual improvements, indicating enhanced food security and reduced reliance on negative coping mechanisms.
3. **Water Access:** The project significantly improved households' access to at least 15 litres of water per person per day, with 58% of treatment households reporting adequate access compared to 32% of control households.
4. **Protection:** The project positively impacted safety feelings and perceptions of safe assistance delivery.

Machine Learning Analysis: The Random Forest model identified key drivers of vulnerability, highlighting the importance of water location, water access, food security (FCS), coping strategies (CSI), safe assistance, body soap access, presence of unaccompanied children, and phase participation.

Gender Analysis: Initially, the intervention benefited women more, but differences in food security outcomes between genders diminished over time. Women faced significant challenges during the early intervention period, with higher instability in food consumption.

Livelihood Coping Strategy Index (LCSI): Significant differences were found in the adoption of various coping strategies between control and treatment groups, indicating a shift from severe to less severe coping mechanisms due to the project intervention. However, a sizable portion of households still relied on emergency coping strategies, highlighting ongoing challenges.

Water Access Analysis:

1. Key Findings:

- Households in the treatment group reported significantly better water access compared to control households (58% vs. 32%).
- The Random Forest model identified critical factors influencing water access, such as proximity to water sources, ability to meet basic water needs, fetching time, availability of soap, and geographic location.

2. Significant Covariates:

- **Water Source Location (wash.location):** Proximity to water sources is significantly associated with improved water access.
- **Basic Water Needs (wash.basic_needs):** Households meeting their basic water needs had better access.
- **Fetching Time (wash.fetchingtime):** Shorter fetching times correlated with better water access.
- **Availability of Soap (wash.nfi_soap):** Availability of soap showed mixed but positive effects on water access.

Protection Analysis:

1. Key Findings:

- The Random Forest model identified significant predictors of safety feelings, including geographic location, quality of protection and community relationships, shelter conditions, health and service availability, and livelihood and market conditions.
- The DiD model indicated significant improvement in safety feelings during the first to third years of implementation.

2. Significant Covariates:

- **Latitude (lat):** Positive and significant association with safety feelings, suggesting that households located further north report better safety feelings.
- **GBV Response (health.sgbv_response):** Better GBV response mechanisms are associated with improved safety perceptions.

Implications for Project Objectives and Future Interventions

The SAHARA project has demonstrated substantial positive impacts across multiple dimensions of household well-being, achieving significant improvements in food security, water access, and protection. These results underscore the importance of comprehensive, multi-faceted interventions tailored to address the specific needs and vulnerabilities of the target populations.

The observed improvements in key indicators such as the Food Consumption Score (FCS), Coping Strategy Index (CSI), and access to water highlight the effectiveness of the project's strategies. The significant reductions in vulnerability scores and enhanced safety perceptions indicate that the project has successfully created a more secure and resilient environment for the affected communities.

Gender Considerations: The gender-specific analysis revealed that women initially benefited more from the interventions, suggesting that gender-sensitive approaches were effective. However, the diminishing differences over time highlight the need for sustained and targeted support for women, particularly during early intervention stages and in response to external shocks.

Livelihood and Coping Strategies: The findings on the Livelihood Coping Strategy Index (LCSI) suggest a positive shift from severe to less severe coping mechanisms, although a sizable portion of households still rely on emergency strategies. This indicates the need for continued support to build sustainable livelihoods and further reduce reliance on negative coping mechanisms.

Water Access and Hygiene: The significant improvements in water access and hygiene underscore the importance of infrastructure projects that reduce the distance to water sources and ensure the availability of hygiene supplies. The challenges faced during the initial implementation phases highlight the need for continuous monitoring and adaptation of strategies to ensure sustained improvements.

Protection and Safety: The improvements in safety feelings and perceptions of safe assistance delivery demonstrate the project's positive impact on community protection. Strengthening mechanisms to address gender-based violence and promoting positive interactions between internally displaced persons (IDPs) and host communities are crucial for enhancing overall safety.

The SAHARA project has achieved significant improvements in food security, water access, and protection in the communities it serves. The findings highlight the importance of early and sustained intervention, comprehensive strategies addressing water and sanitation, and tailored support for vulnerable subgroups. Continued efforts to refine and optimise interventions based on ongoing monitoring and qualitative insights will be crucial for sustaining and enhancing the project's impact on community well-being.

Recommendations for Future Work:

1. Early and Continuous Intervention:
 - Prioritise early and sustained intervention in vulnerable communities to maintain and enhance positive impacts.
2. Enhance Water and Sanitation Infrastructure:
 - Reduce the distance between households and water sources through infrastructure projects.
 - Ensure availability of hygiene supplies, such as soap, to improve water access.
 - Address basic water needs of households to ensure sufficient water for daily use.

- Implement measures to reduce water fetching times, such as increasing the number of water points and improving transportation means.
- 3. Support Food Security Initiatives:
 - Expand investment in food security interventions, such as food assistance programmes, livelihood support, and nutrition education.
- 4. Address Protection Concerns:
 - Develop comprehensive strategies to address safety concerns and improve perceptions of safe assistance delivery.
 - Strengthen mechanisms addressing gender-based violence to boost community safety feelings.
 - Promote positive interactions and relations between IDPs and host communities.
 - Enhance shelter quality and health service availability to contribute to better safety feelings.
- 5. Focus on Vulnerable Subgroups:
 - Tailor interventions to meet the unique needs of households with unaccompanied children and other vulnerable subgroups.
- 6. Targeted Gender-Specific Interventions:
 - Implement interventions addressing unique challenges faced by people in accessing food, particularly during early project stages and in response to external shocks.
- 7. Continuous Monitoring and Evaluation:
 - Regularly monitor and evaluate the impact of different implementation phases on key outcomes to optimise strategies and interventions.
 - The observational nature of the study limits causal inference. Treatment and control assignment based on project coverage at the village level—where households in covered villages are considered treatment households even if they may not have directly benefited—can introduce biases. To enhance causal inference, experimental designs should be considered where feasible, ideally integrating data collection with activity implementation. For example, using forms to register output delivery and capture beneficiary feedback in real-time (as demonstrated in tools like <https://movimentar.shinyapps.io/faith4peace>) can enable more robust and timely evaluations.
 - Longitudinal studies are needed to provide deeper insights into the sustained effects of the project.
- 8. Knowledge Sharing and Dissemination:
 - Share findings and insights with relevant stakeholders to promote evidence-based decision-making in the humanitarian sector.

The significant improvements observed in the SAHARA project underscore the potential for well-designed, comprehensive interventions to bring about meaningful change. Stakeholders, including project managers, donors, and policymakers, are encouraged to invest in similar projects, adopt evidence-based practices, and prioritise the well-being of vulnerable populations. In conclusion, the SAHARA project has made significant strides in improving the lives of the communities it serves. The comprehensive approach, combining food security, water access, protection, and coping strategies, has proven effective in creating a more secure and resilient environment. Future efforts should focus on sustaining and enhancing these positive impacts, ensuring that the benefits reach the most vulnerable populations and addressing any emerging challenges through adaptive and responsive strategies.

6 References

- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- Berge, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: The R package FENmlm. *CREA Discussion Papers*.
- Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3), 1–25. <https://doi.org/10.18637/jss.v040.i03>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1–20. <https://doi.org/10.18637/jss.v021.i12>
- Rinker, T. W. (2018). *textclean: Text cleaning tools (Version 0.9.3) [Computer software]*. Buffalo, New York. Retrieved from <https://github.com/trinker/textclean>