The sub Saharan Journal of Social Sciences and Humanities
(SJSSH)
ISSN: 2619-8894 (Online), 2619- 8851 (Print)

What Drives or Constrains Legume Production in Semi-Arid Smallholder Systems? Unpacking Socio-Economic Determinants, Yield Variability, and Decadal Trends in Mkalama District, Tanzania

[1]Supira S. Kaalai, Edwin E. Ngowi[2], Ng'winamila D. Kasongi[3], Frederick Baijukya[4] and Valery S. Sonola[5]

[1]Department of Policy, Planning and Management, Sokoine University of Agriculture, Morogoro, Tanzania. Email: kalaisupira@gmail.com
[2]Department of Development and Strategic Studies, Sokoine University of Agriculture, Morogoro, Tanzania. Email: edwin.ngowi@sua.ac.tz
[3]Department of Geography and Environmental Studies, The University of Dodoma, Dodoma, Tanzania. Email: nkasongi@gmail.com
[4]International Institute of Tropical Agriculture (IITA), Dar es Salaam, Tanzania. Email: F.Baijukya@cgiar.org
[5]Livestock Training Agency, Buhuri Campus, P. O. Box 1483, Tanga, Tanzania. Email: vssonola@gmail.com

Received: October 2, 2025; **Accepted:** March 21, 2026; **Published:** March 31, 2026

Abstract: *Despite the recognized nutritional, agronomic, and economic roles of grain legumes in sub-Saharan Africa, smallholder yields in semi-arid regions remain persistently low and unstable, with limited empirical understanding of the underlying socio-economic drivers. This study investigated the levels, variability, and ten-year production trends of legumes among smallholder farmers in Mkalama District, Tanzania, and systematically examined the influence of key socio-economic factors on output. A cross-sectional survey of 138 legume-growing households was conducted using structured questionnaires, and data were analyzed using descriptive statistics, one-way ANOVA, and chi-square tests. Results revealed a mean total legume production of 784.52 kg per farmer (SD = 1489.35 kg), indicating extreme variability. Beans dominated as the primary crop (84.06% of farmers), while groundnuts and green gram each accounted for one-third. Over the past decade, production trends were highly unstable: 48.55% of farmers reported fluctuating yields, 35.51% reported declines, and only 2.90% reported increases. One-way ANOVA demonstrated that annual income and farm size significantly influenced legume output ($p < 0.001$). Farmers earning above 2,000,000 TZS produced on average 2678.08 kg compared to 234.13 kg for those earning below 200,000 TZS. Similarly, farmers with >6 acres produced 1477.35 kg, more than five times the output of those with 1–3 acres. In contrast, education level showed no significant effect on production, and chi-square tests found no significant associations between demographic variables (sex, education, income) and production trends ($p > 0.90$). These findings provide two key contributions: empirically, they demonstrate that resource endowments, not demographic characteristics, are the primary drivers of legume productivity in semi-arid smallholder systems; and for policy, they highlight the ineffectiveness of undifferentiated demographic targeting. We recommend prioritized interventions in climate-resilient technologies, equitable input subsidies, market linkages, and support for resource-poor farmers to reduce yield variability and stabilize legume production in semi-arid zones.*

Keywords: *Legume production variability, Socio-economic determinants, Smallholder farmers, Semi-arid agriculture, Tanzania*

1.0 Background Information

Agriculture remains the backbone of Tanzania's economy, contributing approximately 28.7% to the national Gross Domestic Product (GDP) and employing nearly 65% of the labour force, with smallholder farmers operating on plots averaging less than two hectares accounting for over 85% of agricultural output (Mpogole, Dimoso, & Mayaya, 2020; National Bureau of Statistics [NBS], 2021). Within this agrarian economy, grain legumes, including common bean (*Phaseolus vulgaris*), groundnut (*Arachis hypogaea*), pigeon pea (*Cajanus cajan*), cowpea (*Vigna unguiculata*), chickpea (*Cicer arietinum*), and soybean (*Glycine max*), occupy a strategic position, covering over 1.5 million hectares and producing approximately 1.1 million metric tons annually (Abate, 2020; Ministry of Agriculture [MoA], 2022). Legumes contribute dually to smallholder livelihoods: they serve as primary protein sources in cereal-dominated diets and generate cash income through local and regional markets (Vanlauwe *et al.*, 2019; Kinyua *et al.*, 2023).

The agronomic significance of legumes extends beyond direct consumption. Their capacity for biological nitrogen fixation (BNF) enhances soil fertility, reduces synthetic fertilizer dependence, and improves cropping system resilience—benefits particularly critical in resource-poor, low-input farming systems characteristic of sub-Saharan Africa (Stagnari, Maggio, Galieni, & Pisante, 2017; Peoples *et al.,* 2019). In Tanzania's semi-arid zones, where soil degradation and declining fertility are pervasive, legume integration into cereal-based systems offers ecological and economic synergies that support sustainable intensification (Snapp et al., 2018; Franke, van den Brand, & Giller, 2019).

Despite these recognized benefits, legume productivity in Tanzania remains substantially below agronomic potential. National average yields for common bean range between 500–800 kg ha^{-1}, compared to potential yields exceeding 1,500 kg ha^{-1} under improved management (Venance, Mshenga, & Birachi, 2016; Katungi, Magreta, Letaa, & Chirwa, 2017). Similarly, groundnut yields average 700–900 kg ha^{-1} against a potential of 2,000–2,500 kg ha^{-1} (Okello, Monyo, Deom, & Gichuki, 2018). This persistent yield gap reflects multiple interacting constraints: limited access to improved seeds and inputs, poor soil fertility, pest and disease pressure, erratic rainfall, weak extension services, and market inefficiencies (Nord *et al.,* 2022; Lugamara *et al.,* 2021; Mponela, Tamene, Ndengu, & Magreta, 2020).

In Mkalama District, located in Singida Region's semi-arid corridor (annual rainfall 500–800 mm), legumes, particularly beans, groundnuts, and green gram, constitute key cash and food crops. However, district-level agricultural reports indicate substantial production variability and declining trends over recent years (Mkalama District Council, 2021). Farmers report erratic yields, frequent crop failure, and limited capacity to respond to climate shocks, suggesting a complex interplay of resource constraints, institutional gaps, and environmental stressors (Simel, 2026). Yet, systematic empirical evidence quantifying legume production levels, trends, and their socio-economic determinants in this under-researched locality remains scarce.

Previous studies in Tanzania have examined legume adoption dynamics, technology diffusion, and agronomic performance. Lugamara *et al.* (2021) identified household size, farm area, extension access, and group membership as determinants of improved legume technology adoption across multiple districts. Nord *et al.* (2022) documented critical disconnects between extension messaging and farmer practice in maize-legume systems, noting that recommended practices often misalign with local resource realities. Kinyua *et al.* (2023) demonstrated that legume-legume and cereal-legume intercropping configurations can improve productivity and resilience but emphasized the need for context-specific adaptation. However, few investigations have systematically quantified socio-economic determinants underlying legume output levels, production trends, and yield variability among farmers in semi-arid zones, and none have specifically addressed Mkalama District.

Furthermore, evidence from elsewhere in East Africa highlights the importance of resource endowment in shaping agricultural outcomes. Mugi (2022) found that disparities in input access, irrigation, and market linkages drive legume yield variability in northern Tanzania. In Ethiopia, Abebe, Deressa, and Dechassa (2022) reported that targeted extension and input subsidies reduced productivity gaps between resource-rich and resource-poor farmers. Similarly, Franke *et al.* (2019) demonstrated that socio-economic factors, particularly access to labour, capital, and land, explain more variation in climbing bean productivity than biophysical conditions in Rwanda. These findings suggest that socio-economic determinants may be as consequential as agronomic factors in shaping legume production outcomes.

Understanding these dynamics is particularly urgent given the mounting challenges facing Tanzania's agricultural sector: climate variability (IPCC, 2022), persistent food insecurity (FAO, 2021), rural poverty (World Bank, 2020), and structural market limitations (Mmbando & Baiyegunhi, 2021). Legumes, with their nutritional, agronomic, and economic attributes, offer a potentially transformative lever for sustainable development, but only if production constraints are systematically diagnosed and addressed. This requires moving beyond aggregate production statistics to examine the distribution, determinants, and dynamics of legume productivity across heterogeneous smallholder populations.

Therefore, this study aims to: (1) quantify production volumes and variability of key legumes; beans, groundnuts, and green gram; among farmers in Mkalama District; (2) evaluate production trends over the past ten years, identifying patterns of increase, decline, or fluctuation; (3) analyse the influence of socio-economic factors (income, farm size, education level) on legume production levels; and (4) determine associations between demographic characteristics (sex, education, income) and legume production trends over time. Therefore, by generating empirical evidence on these dimensions, the study seeks to inform targeted policy and programmatic interventions that can enhance legume productivity, stabilize yields, and improve smallholder livelihoods in semi-arid Tanzania.

2.0 Theoretical Framework

This study is anchored on an integrated theoretical framework that draws from three complementary theoretical perspectives: (1) the Sustainable Livelihoods Framework (SLF), (2) the Theory of Resource-Based View (RBV), and (3) the Farm Household Model. The integration of these theories provide a strong analytical lens for understanding the socio-economic determinants, production trends, and variability of legume yields among smallholder farmers in semi-arid Mkalama District, Tanzania.

2.1 Sustainable Livelihoods Framework (SLF)

The Sustainable Livelihoods Framework, developed by the Department for International Development (DFID, 1999) and extensively refined by Scoones (2009) and Morse and McNamara (2013), provides a holistic approach to understanding how households combine various capital assets to pursue livelihood outcomes. The framework posits that households possess five types of capital assets: human capital (skills, knowledge, labour, education), social capital (networks, group membership, trust), natural capital (land, water, soil quality, biodiversity), physical capital (infrastructure, tools, equipment, storage facilities), and financial capital (savings, credit, income, remittances). These assets are shaped by the vulnerability context (shocks, trends, seasonality) and transforming structures and processes (institutions, policies, governance), which together influence livelihood strategies and outcomes (Scoones, 2015; Morse, McNamara, & Acholo, 2018).

In the context of legume production among smallholder farmers, the SLF is particularly relevant because it recognises that agricultural productivity is not solely determined by biophysical factors or individual farmer characteristics, but by the constellation of assets farmers can mobilise (Ellis, 2016; Glover, Sumberg, & Andersson, 2019). For instance, a farmer's ability to purchase improved legume seeds, fertilisers, and pesticides depends on financial capital; access to extension services and training depends on human and social capital; and the area and quality of land allocated to legumes depends on natural capital. Furthermore, the vulnerability context, including climate variability, market price fluctuations, and pest outbreaks, directly shapes production trends and yield stability (Tittonell, 2020; Vanlauwe *et al.,* 2019). This study therefore operationalises the SLF by examining how different forms of capital (income as financial capital, farm size as natural capital, education as human capital, and demographic characteristics as proxies for social capital) influence legume production levels and trends.

2.2 Resource-Based View (RBV) Theory

The Resource-Based View, originating from strategic management literature (Barney, 1991; Peteraf, 1993), offers a complementary theoretical lens by focusing on how heterogeneous resource endowments generate differential performance outcomes. The central tenet of RBV is that firms (or, in this adaptation, farm households) achieve competitive advantage not merely through industry positioning but through ownership of valuable, rare, inimitable, and non-substitutable (VRIN) resources (Barney, Ketchen, & Wright, 2021). In smallholder agriculture, this translates to understanding how differential access to productive resources; land, labour, capital, and knowledge; creates persistent productivity gaps between farmer groups (Wiggins, 2014; Giller *et al.,* 2021).

Applied to legume production, RBV predicts that farmers with superior resource endowments (larger landholdings, higher income, better access to information) will achieve higher and more stable yields because they can invest in quality inputs, adopt improved technologies, and buffer against shocks (Muyanga & Jayne, 2019; Ren, Liu, Li, & Zhao, 2019). Conversely, resource-poor farmers face binding constraints that trap them in low-productivity equilibria (Barrett & Bevis, 2015). This theoretical expectation aligns with the study's ANOVA results, which show that farmers earning above 2,000,000 TZS produce significantly more legumes (mean 2678.08 kg) than those earning below 200,000 TZS (mean 234.13 kg), and farmers with over six acres produce more than five times the output of those with one to three acres. The RBV thus provides theoretical grounding for examining income and farm size as critical determinants of legume production variability.

2.3 Farm Household Model

The Farm Household Model, rooted in agricultural economics (Singh, Squire, & Strauss, 1986; De Janvry, Fafchamps, & Sadoulet, 1991), conceptualises smallholder farmers as simultaneously producers and consumers who make household decisions under multiple constraints, resource constraints (land, labour, capital), market constraints (transaction costs, price risk, imperfect markets), and risk constraints (production risk, price risk, health shocks) (Bellemare, 2018; Dillon & Barrett, 2017). This model recognises that smallholder production decisions, including legume cultivation, are not purely profit-maximising but are shaped by household consumption needs, risk preferences, and market access limitations (Mmbando & Baiyegunhi, 2021; Jayne, Fox, Fuglie, & Adelaja, 2021).

For legume production in semi-arid Tanzania, the Farm Household Model explains why a majority of farmers (74.64% in this study) grow legumes for both household consumption and income generation, a "*food-cash*" strategy that buffers against market failures and food insecurity (Bonuedi, 2021; IFAD, 2019). The model also illuminates why education alone may not significantly influence legume output: in settings with imperfect markets and limited agriculture-specific training, general education does not necessarily translate into productivity-enhancing practices (Paltasingh & Goyari, 2018; Evenson & Gollin, 2003). Instead, binding resource and risk constraints may override any potential human capital effects, a finding consistent with the study's chi-square results showing no significant associations between demographic variables and production trends.

ISSN: 2619-8894 (Online), 2619- 8851 (Print)

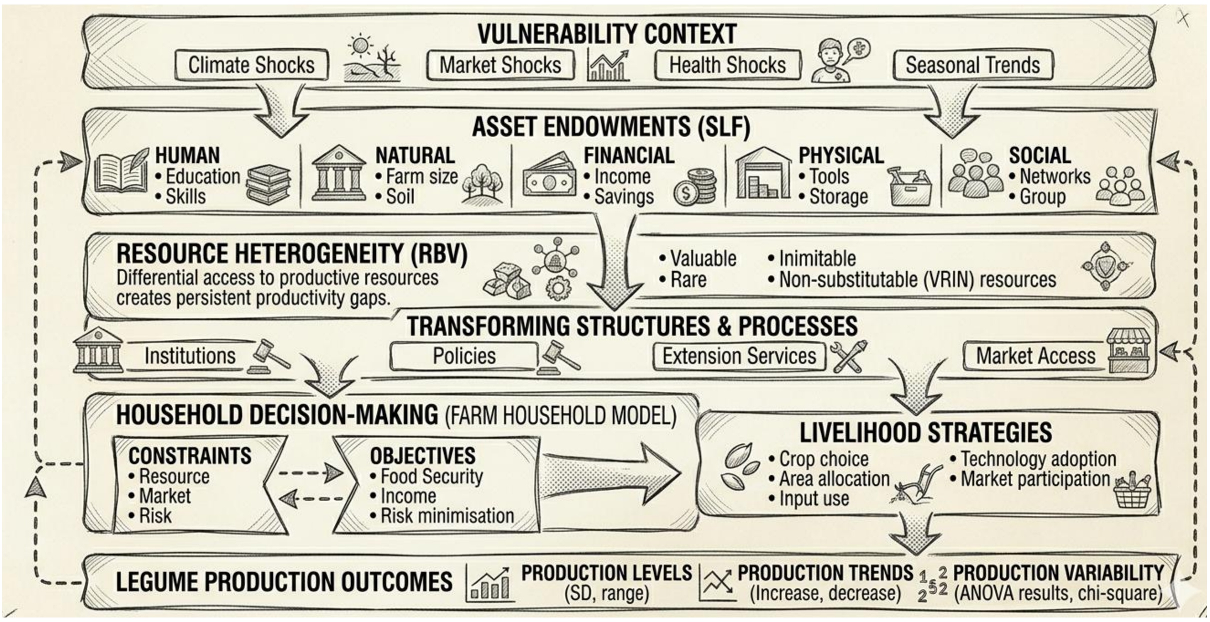
2.4 Integrated Theoretical Framework for Legume Production Among Smallholder Farmers

The integration of the Sustainable Livelihoods Framework, Resource-Based View, and Farm Household Model into a unified theoretical framework (Figure 1) provides a comprehensive analytical structure for this study. The framework posits that legume production outcomes (production levels, trends, and variability) are determined by the interaction of three interconnected domains: (1) asset endowments (human, natural, financial, physical, and social capital as per SLF), (2) resource heterogeneity (differential access to VRIN resources as per RBV), and (3) household decision-making under multiple constraints (production, consumption, risk, and market constraints as per Farm Household Model). These domains are embedded within a vulnerability context (climate variability, market volatility, policy environment) and are mediated by transforming structures (institutions, extension services, market access). The framework guides this study's empirical investigation by directing attention to how specific socio-economic factors, particularly income (financial capital) and farm size (natural capital), influence legume production, while also explaining why demographic characteristics (sex, education) show no significant associations with production trends.

2.5 Application of the Framework to the Present Study

The integrated theoretical framework directly informs the study's empirical design and hypothesis testing. First, the framework guides the selection of socio-economic variables: income (financial capital) and farm size (natural capital) are hypothesised as strong determinants of legume production

Figure 1: Integrated Theoretical Framework for Analysing Legume Production Among Smallholder Farmers in Semi-Arid Tanzania



Source: Authors' construction, adapted from DFID (1999), Barney (1991), and Singh, Squire, and Strauss (1986)

The integrated theoretical framework consists of seven hierarchical components:

***Component 1: Vulnerability Context** i.e., External shocks and trends (climate, market, health, seasonal) that farmers cannot control but must respond to, shaping production risk and yield stability.*

***Component 2: Asset Endowments (SLF)** i.e., The five capital assets (human, natural, financial, physical, social) that farmers possess. This study operationalises these as education and labour (human), farm size and soil fertility (natural), income and credit access (financial), tools and storage (physical), and group membership (social).*

***Component 3: Resource Heterogeneity (RBV)** i.e., Differential access to VRIN resources creates persistent productivity gaps; farmers with superior endowments achieve higher and more stable legume yields.*

***Component 4: Transforming Structures & Processes** i.e., Institutions, policies, extension services, and market access that mediate how assets translate into production outcomes.*

***Component 5: Household Decision-Making (Farm Household Model)** i.e., Farmers make legume production decisions under multiple constraints (resource, market, risk) to achieve multiple objectives (food security, income, risk minimisation).*

***Component 6: Livelihood Strategies** i.e., Specific legume production choices including crop selection, area allocation, input use, technology adoption, and market orientation.*

***Component 7: Legume Production Outcomes** i.e., The dependent variables: production levels (kg per farmer), production trends (decadal patterns), and production variability (yield dispersion).*

***Feedback Loops:** Production outcomes feed back to affect future asset endowments, creating dynamic pathways out of or into poverty traps.*

ISSN: 2619-8894 (Online), 2619- 8851 (Print)

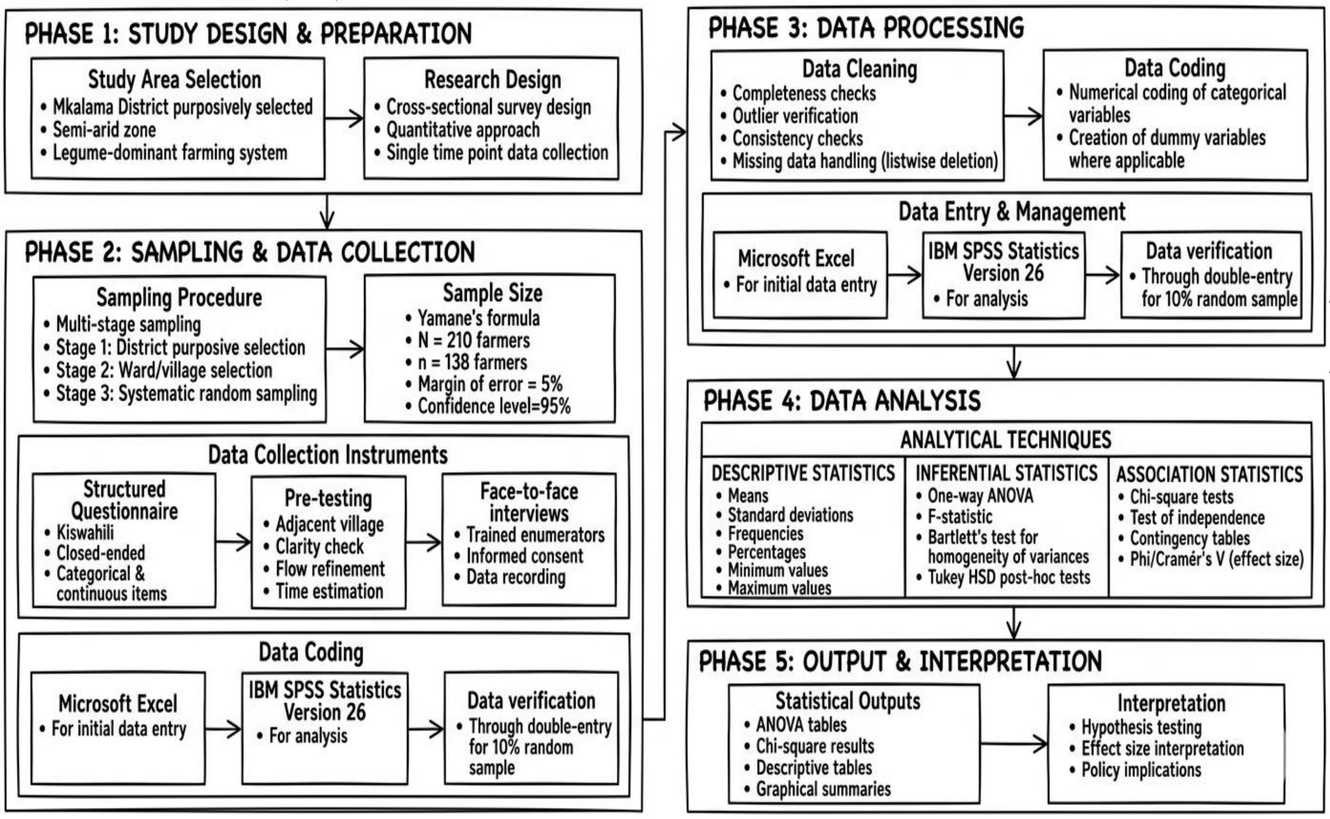
levels (supported by ANOVA results, $p < 0.001$), while education (human capital) may show weaker effects due to the absence of agriculture-specific training and binding resource constraints (supported by non-significant post-hoc comparisons). Second, the framework explains the high production variability (SD = 1489.35 kg) and predominantly fluctuating/declining trends (48.55% fluctuation, 35.51% decline) as outcomes of vulnerability context shocks (climate variability, market volatility) interacting with heterogeneous asset endowments. Third, the framework accounts for the absence of significant associations between demographic factors and production trends (chi-square $p > 0.90$) by recognising that structural constraints; market access, climate shocks, input availability; may exert more uniform influences across demographic groups than hypothesised. The framework thus provides theoretical coherence to the empirical findings and offers a basis for policy recommendations targeting asset building (particularly financial and natural capital), vulnerability reduction, and transforming structures (extension, market access).

3.0 Methodology

Figure 2 presents a schematic overview of the methodological framework guiding this study, while Table 1 provides a systematic alignment between research objectives, data requirements, analytical methods, and expected outputs. The methodology section is organised into seven subsections: study area description, research design, sampling procedure and sample size determination, data collection instruments and procedures, data processing and management, analytical framework linking methods to specific objectives, and ethical assurances.

The methodological framework (Figure 2) comprises five sequential phases. Phase 1 establishes the study design, including purposive selection of Mkalama District based on its semi-arid conditions and legume-dominant farming systems, and adoption of a cross-sectional survey design. Phase 2 details the multi-stage sampling procedure and sample size determination using Yamane's formula, followed by data collection using pre-tested structured questionnaires administered through face-to-face interviews. Phase 3 encompasses data processing activities including cleaning, coding, and management using SPSS software. Phase 4 specifies the analytical techniques employed: descriptive statistics for production quantification, one-way ANOVA for comparing means across socio-economic categories, and chi-square tests for examining associations between demographic variables and production trends. Phase 5 presents the statistical outputs and their interpretation in the context of the study objectives and policy implications.

Figure 2: Methodological Framework for Assessing Legume Production Determinants and Trends in Mkalama District, Tanzania



Source: Authors' construction based on methodological approaches in agricultural household surveys (Grosh & Glewwe, 2000; Carletto, Zezza, & Banerjee, 2013)

Table 1: Alignment of Research Objectives with Data Requirements, Analytical Methods, and Outputs

Objective	Data Required	Data Type	Analytical Method	Statistical Test / Output	Interpretation Focus
Objective 1: Quantify production volumes and variability of key legumes (beans, groundnuts, green gram)	• Quantity of beans produced (kg) in most recent season • Quantity of groundnuts produced (kg) • Quantity of green gram produced (kg) • Total legume production (kg)	Continuous (ratio scale)	Descriptive statistics	• Mean • Standard deviation • Minimum value • Maximum value • Coefficient of variation (CV)	Magnitude and dispersion of legume output across farmers; identification of crop-specific production patterns
Objective 2: Evaluate production trends over past ten years	• Farmer recall of production changes (past decade) • Categorical responses: increased, decreased, fluctuated, stayed same, other	Categorical (nominal)	Frequency distribution analysis	• Frequency counts • Percentage distribution • Modal trend category	Prevalence of yield instability; direction of decadal production trajectories
Objective 3: Analyse influence of socio-economic factors (income, farm size, education) on legume production levels	• Annual income category (TZS) • Farm size category (acres) • Education level (none, primary, secondary, college) • Total legume production (kg)	Independent: categorical (ordinal/nominal) Dependent: continuous	One-way Analysis of Variance (ANOVA)	• F-statistic • p-value • Bartlett's test (homogeneity) • Tukey HSD post-hoc (pairwise differences)	Whether income, farm size, and education significantly affect production; direction and magnitude of differences between groups
Objective 4: Determine associations between demographic characteristics and production trends	• Sex (male/female) • Education level • Income category • Production trend category (from Objective 2)	Both independent and dependent: categorical	Chi-square test of independence	• χ^2 statistic • Degrees of freedom • p-value • Cramér's V (effect size)	Whether demographic variables are systematically associated with production trend patterns

Source: Authors' construction adapted from analytical frameworks in agricultural economics (Greene, 2018; Wooldridge, 2016; Gujarati & Porter, 2009)

Table 1 provides a systematic mapping of the four research objectives to their corresponding data requirements, analytical methods, statistical outputs, and interpretation foci. The alignment ensures methodological coherence and transparency, demonstrating how each objective is operationalised through specific analytical procedures (Onwuegbuzie & Leech, 2005). For Objective 1, descriptive statistics quantify production volumes and variability. For Objective 2, frequency distributions summarise decadal production trends. For Objective 3, one-way ANOVA tests mean production differences across socio-economic categories, with Tukey HSD post-hoc tests identifying which specific groups differ significantly (Field, 2018). For Objective 4, chi-square tests examine whether demographic characteristics are independent of production trend categories (Agresti, 2019). The table is cited in the methodology section to guide readers through the analytical logic.

3.1 Study Area

The study was conducted in Mkalama District, located in Singida Region, central Tanzania (approximately latitude 4°30' to 5°00' S and longitude 34°30' to 35°00' E). The district is characterised by semi-arid climatic conditions with unimodal rainfall averaging 500–800 mm per annum, predominantly falling between November and April, followed by a prolonged dry season of five to six months (Mkalama District Council, 2021). Mean annual temperatures range from 20°C to 28°C, with high evapotranspiration rates typical of central Tanzania's plateau ecosystems (Lyimo, 2019). Soils are predominantly sandy loams and clay loams with low organic matter content (less than 1.5%) and moderate to high phosphorus fixation, which constrains legume nodulation and biological nitrogen fixation unless specific rhizobia strains are present (Semu, Singh, & Msaky, 2018).

The agricultural economy relies heavily on smallholder mixed farming systems, with households cultivating cereals (maize, sorghum, millet) and legumes (beans, groundnuts, green gram, pigeon pea) as both food and cash crops (Mpogole, Dimoso, & Mayaya, 2020). Livestock keeping (cattle, goats, sheep) is integrated into farming systems as an asset accumulation and risk diversification strategy (Mtengeti *et al.,* 2015). Mkalama District was purposively selected for this study because of (i) its known dependence on legume production for household food security and income generation, (ii) reports from district agricultural officers of inconsistent production trends and declining yields over recent years, (iii) its representation of semi-arid conditions prevalent across much of central and northern Tanzania, and (iv) the absence of prior systematic research on legume production determinants in the district (Simel, 2026).

3.2 Research Design

A cross-sectional survey design was adopted to collect quantitative data at a single point in time from multiple households. This design was appropriate for capturing production volumes, historical trends (using recall data), and socio-economic characteristics, and for statistically testing associations between variables (Kothari, 2014; Creswell & Creswell, 2018). Cross-sectional designs are widely used in agricultural household surveys because they allow efficient estimation of population parameters and hypothesis testing within resource and time constraints, although they cannot establish causal ordering (De Vaus, 2013). To minimise recall bias for historical production trends, the questionnaire used specific reference periods and categorical response options (increased, decreased, fluctuated, stayed the same) rather than requiring precise annual yield estimates (Tadesse, Shiferaw, & Erenstein, 2015).

3.3 Sampling Procedure

A multi-stage sampling technique was used following procedures recommended by Cochran (1977) and Thompson (2012) for household surveys in agricultural contexts. The sampling proceeded through three stages:

Stage 1 (District selection): Mkalama District was purposively selected based on its semi-arid agroecology, predominance of legume cultivation, reported production instability, and lack of prior research attention. Purposive selection is appropriate when the study requires specific characteristics not randomly distributed across the population (Patton, 2015).

Stage 2 (Ward and village selection): From the district's administrative wards, three wards with high prevalence of legume farming were identified in consultation with district agricultural officers and village extension agents. Within each selected ward, two villages with active legume farmer populations were purposively selected, yielding six villages total. This two-stage purposive selection ensured the sample captured farmers with relevant production experience (Bernard, 2017).

Stage 3 (Household selection): Systematic random sampling was used to select respondents from village household registers maintained by village executive officers. A sampling interval (k = N/n_village) was calculated for each village, and starting points were randomly selected. Systematic random sampling provides probability-based selection while being logistically feasible in rural settings where complete sampling frames exist (Lohr, 2019).

3.4 Sample Size Determination

The sample size was determined using Yamane's (1967) formula, which is widely used in agricultural and social science research for finite populations (Israel, 2013; Bartlett, Kotrlik, & Higgins, 2001). The formula is expressed as:

$$n = \frac{N}{1 + N(e)^2}$$

Where:

- n = required sample size
- N = total number of legume farmers in the sampling frame (target population)
- e = desired margin of error (0.05 for 95% confidence level)

The target population comprised 210 legume farmers listed in the six village household registers who had cultivated legumes in the most recent agricultural season. Applying Yamane's formula:

$$n = \frac{210}{1 + 210(0.05)^2} = \frac{210}{1 + 210(0.0025)} = \frac{210}{1 + 0.525} = \frac{210}{1.525} = 137.7 \approx 138$$

Thus, 138 legume farmers were surveyed. This sample size exceeds the minimum required for one-way ANOVA with four groups (assuming medium effect size, $\alpha = 0.05$, power = 0.80), which would require approximately 128 observations (Cohen, 1988; Faul, Erdfelder, Buchner, & Lang, 2009). The achieved sample of 138 provides adequate statistical power for the planned ANOVA and chi-square analyses.

3.5 Data Collection Instruments and Procedures

Data were collected using structured questionnaires administered through face-to-face interviews in Kiswahili, the local language widely spoken in Singida Region (Mkalama District Council, 2021). The questionnaire was developed based on a review of agricultural household survey instruments used in similar Tanzanian contexts

(Venance, Mshenga, & Birachi, 2016; Lugamara *et al.,* 2021; Nord *et al.,* 2022) and included four sections:

Section A: Household demographics: Sex, age, education level, household size, primary occupation.

Section B: Socio-economic characteristics: Annual income category (below 200,000 TZS; 200,000–500,000 TZS; 500,000–1,000,000 TZS; 1,000,000–2,000,000 TZS; above 2,000,000 TZS), farm size category (1–3 acres; 4–6 acres; more than 6 acres), land tenure status, and main purpose for legume cultivation (household consumption only; income generation only; both consumption and income; other).

Section C: Legume production data: Quantities (in kilograms) of beans, groundnuts, and green gram produced in the most recent completed agricultural season, using standard local measurement units (debe, gunia) converted to kilograms based on crop-specific conversion factors established by the Tanzania Bureau of Standards (TBS, 2015).

Section D: Historical production trends: Farmers' recall of changes in legume production over the past ten years (2013–2023), categorised as: increased significantly, increased slightly, stayed the same, fluctuated/varied, decreased slightly, decreased significantly, or other patterns. Categorical recall methods reduce recall bias compared to requesting precise annual yield estimates (Beegle, Carletto, & Himelein, 2012).

The questionnaire was pre-tested in a nearby village (Ibaga village, Mkalama District) not included in the final sample. Pre-testing with 15 farmers assessed clarity, question flow, cultural appropriateness, and average administration time (45–60 minutes). Revisions were made to rephrase ambiguous questions, adjust conversion factors based on local measurement practices, and improve response category exhaustiveness (Dillman, Smyth, & Christian, 2014).

Six enumerators (three male, three female) with prior experience in agricultural surveys and fluency in Kiswahili were recruited and trained for two days on the study objectives, questionnaire administration, ethical protocols (informed consent, confidentiality), and accurate recording of production quantities using standardised conversion tables. Fieldwork was conducted during July–August 2023, immediately following the main harvest season (May–June), to maximise recall accuracy for production quantities (Grosh & Glewwe, 2000).

3.6 Data Processing and Management

Data were processed following established protocols for agricultural household survey data (Carletto, Zezza, & Banerjee, 2013). Completed questionnaires were checked daily for completeness and consistency by the field supervisor. Data were entered into Microsoft Excel (version 2019) using a pre-designed template with range and consistency checks to minimise entry errors. For quality assurance, 10% of questionnaires (n = 14) were randomly selected and double-entered; the error rate was below 2%, which is acceptable for survey data (Lohr, 2019). The dataset was then exported to IBM SPSS Statistics (version 26) for cleaning, coding, and analysis. Missing data were handled using listwise deletion, which is appropriate when missingness is less than 5% and appears random (Little & Rubin, 2020).

3.7 Analytical Framework

The analytical framework directly operationalises the four research objectives as presented in Table 1 (cited above). Detailed descriptions of each analytical method follow:

3.7.1 Descriptive Statistics (Objective 1)

Descriptive statistics (mean, standard deviation, minimum, maximum) were computed for each crop (beans, groundnuts, green gram) and total legume production. The coefficient of variation (CV = SD/mean × 100) was calculated to compare relative variability across crops with different mean yields (Gomez & Gomez, 1984). Descriptive statistics provide a quantitative summary of production levels and their dispersion, addressing the magnitude and variability of legume output among farmers.

3.7.2 Frequency Distributions (Objective 2)

Frequency distributions and percentages were used to summarise farmers' reported production trends over the past decade. This non-parametric approach is appropriate for categorical data and identifies the prevalence of different trend patterns (increase, decrease, fluctuation, stability) in the population (Agresti, 2019).

3.7.3 One-way Analysis of Variance (Objective 3)

One-way Analysis of Variance (ANOVA) tested differences in mean total legume production across categories of income, farm size, and education level. ANOVA is appropriate when the independent variable is categorical (with two or more groups) and the dependent variable is continuous (Field, 2018). The null hypothesis was that all group means are equal (H_0: $\mu_1 = \mu_2 = ... = \mu_k$), while the alternative hypothesis was that at least one group mean differs significantly (H_1: not all μ_i are equal). The F-statistic was computed as:

$$F = \frac{MS_{between}}{MS_{within}} = \frac{\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2/(N-k)}$$

Where:

- $MS_{between}$ = mean square between groups
- MS_{within} = mean square within groups (error)
- k = number of groups
- n_i = sample size of group i
- N = total sample size

Bartlett's test assessed the assumption of homogeneity of variances across groups. Where Bartlett's test indicated violation of homogeneity ($p < 0.05$), Welch's ANOVA was considered as a robust alternative (Field, 2018). Where ANOVA indicated significant overall differences ($p < 0.05$), Tukey's Honestly Significant Difference (HSD) post-hoc tests identified which specific group pairs differed significantly, controlling for family-wise error rate (Tukey, 1949; Abdi & Williams, 2010).

3.7.4 Chi-square Tests of Independence (Objective 4)

Chi-square tests of independence examined relationships between categorical demographic variables (sex, education level, income category) and the categorical production trend variable. The chi-square statistic (χ^2) was computed as:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Where:

- O_{ij} = observed frequency in cell ij
- E_{ij} = expected frequency in cell ij under independence = (row total × column total) / grand total
- r = number of rows
- c = number of columns

The null hypothesis was that the row and column variables are independent (no association), while the alternative hypothesis was that they are associated (Agresti, 2019). Cramér's V was computed as an effect size measure for significant associations, where $V = \sqrt{\chi^2/(n \times \min(r-1, c-1))}$ with values of 0.1, 0.3, and 0.5 interpreted as small, medium, and large effects respectively (Cohen, 1988). All statistical tests were conducted at $\alpha = 0.05$ significance level (two-tailed). Statistical analyses were performed using IBM SPSS Statistics version 26 (IBM Corp., 2019).

3.8 Ethical Considerations

Permission to conduct the study was obtained from the Sokoine University of Agriculture (REF No. *SUA/ADM/R.1/8/1323*) and the President's Office Regional Administration and Local Government (PORALG) prior to commencement of fieldwork. The study complied with Tanzania's national research ethics guidelines as outlined in the Tanzania National Health Research Ethics Guidelines (National Institute for Medical Research, 2015) and principles of the Belmont Report (respect for persons, beneficence, justice) (National Commission for the Protection of Human Subjects, 1979). All participants were briefed on the purpose of the study, the voluntary nature of participation, their right to withdraw at any time without consequence, and how data would be used and stored. Written informed consent was obtained from each participant before data collection. Data were anonymised using unique participant codes rather than names to maintain confidentiality. No personally identifiable information was retained in the final dataset. The study involved minimal risk to participants, as questions focused on agricultural production and household characteristics without sensitive topics.

4.0 RESULTS AND DISCUSSION

This section integrates the presentation of empirical findings with their theoretical and contextual interpretation, following best practices for high-impact journal publication (Makin & Orban de Xivry, 2019; Ghasemi, 2021). The results are organised sequentially according to the four research objectives, with each subsection presenting statistical outputs (descriptive statistics, ANOVA, chi-square tests) followed immediately by discussion situating findings within existing literature and the theoretical framework developed in Section 2.0. All tables are numbered sequentially from Table 2 (continuing from Table 1 presented in the methodology section) and are cited within the text.

4.1 Legume Production Volumes and Variability

Table 2 presents the mean, standard deviation, minimum, and maximum production volumes for total legumes and specific legume types (beans, groundnuts, and green gram) among 138 farmers in Mkalama District. The mean total production was 784.52 kg per farmer, with a strikingly high standard deviation (SD = 1489.35 kg), indicating extreme variability in legume output across the sample. The coefficient of variation (CV = SD/mean × 100 = 190%) confirms that production is highly dispersed, with some farmers producing as little as 15 kg and others producing up to 12,450 kg annually. This level of variability is substantially higher than that reported in similar smallholder legume systems in sub-Saharan Africa. For instance, Mugi (2022) documented CVs ranging from 85% to 120% for legume intercropping systems in northern Tanzania, while Franke et al. (2019) reported CVs of 95–110% for climbing bean productivity in Rwanda. The exceptionally high CV (190%) in Mkalama suggests that production disparities are more extreme in this semi-arid context, likely reflecting heterogeneous access to irrigation, inputs, and climate adaptation strategies.

Beans had the highest average production (386.04 kg) and were the most widely grown crop (n = 111 farmers, 84.06% of the sample, as shown in Table 3). This finding aligns with Mpogole, Dimoso, and Mayaya (2020), who identified common bean as the primary legume in Tanzanian smallholder systems due to its cultural importance in maize-based diets and established market channels. Similarly, Abate (2020) reported that beans account for approximately 60% of legume area in Tanzania, a proportion consistent with the 84% farmer prevalence observed in Mkalama. The mean bean yield (386.04 kg per farmer) must be interpreted with caution because it reflects total production per farmer rather than per hectare yield. However, assuming average bean plot sizes of 0.5–1.0 acres in semi-arid Singida Region (Mkalama District Council, 2021), the implied per-hectare yield (approximately 800–1,200 kg ha^{-1}) falls within the national average range (500–800 kg ha^{-1}) reported by Venance, Mshenga, and Birachi (2016) and Katungi et al. (2017), suggesting that bean productivity in Mkalama is typical of Tanzanian smallholder systems.

Groundnuts (mean 151.43 kg, SD = 142.90 kg) and green gram (mean 1298.41 kg, SD = 2587.09 kg) were each cultivated by approximately one-third of farmers (46 farmers, 33.33% each, Table 3). Green gram showed the largest production variability (SD = 2587.09 kg, CV = 199%), driven by a small number of farmers (n = 5) producing exceptionally high quantities (4,000–12,450 kg). These high-output farmers likely have larger land allocations, better market access, or more favourable microclimates for green gram production. The lower adoption of groundnuts and green gram compared to beans may reflect agroecological constraints (groundnuts require well-drained sandy soils, which are patchily distributed in Mkalama), less-developed value chains, or lower consumer preference (Bogale, 2021; Voss *et al.,* 2024). Vilakazi et al. (2025) similarly found that neglected grain legumes (including green gram) suffer from underinvestment in seed systems and extension despite their nutritional and drought-tolerance advantages.

heterogeneous resource endowments (land, capital, information) generate persistent performance gaps. The finding that green gram production is both low-adoption (33% of farmers) and highly variable (CV = 199%) suggests that green gram may be a “niche crop” suited to specific resource-endowed farmers who can invest in its production and access specialised markets (Ren *et al.,* 2019).

4.2 Primary Legume Types Cultivated

Table 3 presents the distribution of the main legumes grown by sampled farmers. Beans dominated the legume portfolio, cultivated by 84.06% of respondents (116 farmers), reflecting their status as a staple food and cash crop in semi-arid Tanzania. Groundnuts and green gram were each grown by one-third of farmers (33.33%, 46 farmers each). Multiple responses were permitted, as some farmers cultivate more than one legume type (polyculture or sequential cropping).

The dominance of beans aligns with national patterns documented by Nassary (2025), who reported that common bean is the most widely cultivated legume in Tanzania's northern highlands, and by Vilakazi *et al.* (2025), who noted beans' centrality to food and nutrition security across sub-Saharan Africa. The cultural preference for beans in Tanzanian diets, particularly in the staple dish *ugali* served with bean sauce, creates stable demand, encouraging farmers to allocate land to beans even when yields are suboptimal (Mpogole *et al.,* 2020). Additionally, beans have shorter maturity periods (70–100 days) compared to groundnuts (90–120 days) and pigeon pea (180–240 days), making them more compatible with the short unimodal rainy season in semi-arid Mkalama (Abate, 2020).

Table 2: Legume Production Volume and Variability Among Smallholder Farmers (n = 138)

Variable	Observations (n)	Mean (kg)	Standard Deviation (kg)	Minimum (kg)	Maximum (kg)	Coefficient of Variation (%)
Total Production (all legumes)	138	784.52	1489.35	15	12450	190
Beans Production	111	386.04	583.43	25	4200	151
Groundnuts Production	42	151.43	142.90	30	650	94
Green Gram Production	44	1298.41	2587.09	85	12450	199

Note: Coefficient of variation (CV) = (Standard Deviation / Mean) × 100. Observations differ by crop because not all farmers cultivate each legume type. Source: Authors' survey data (2023)

The high production variability documented in Table 2 is consistent with the Sustainable Livelihoods Framework (DFID, 1999; Scoones, 2015) presented in Figure 1, which posits that differential asset endowments (financial, natural, physical capital) produce heterogeneous livelihood outcomes. Farmers with larger landholdings, higher income, and better access to irrigation and inputs (the resource-rich minority) achieve high legume output, while resource-poor farmers remain trapped in low-productivity equilibria. This interpretation aligns with the Resource-Based View (Barney, 1991; Barney, Ketchen, & Wright, 2021), which predicts that

The equal prevalence of groundnuts and green gram (both 33.33%) is noteworthy because national data typically show groundnuts occupying a larger area (approximately 25% of legume area) than green gram (approximately 10%) (Ministry of Agriculture, 2022). The relatively high green gram adoption in Mkalama may reflect the crop's drought tolerance and short maturity (60–75 days), which provides a risk-management strategy in a semi-arid environment with unpredictable rainfall (Okello *et al.,* 2018). Farmers may plant green gram when rainfall onset is delayed or when beans have failed, consistent with the Farm Household

Model's prediction that smallholders diversify crop portfolios to manage production risk (Singh, Squire, & Strauss, 1986; Bellemare, 2018). However, the lower absolute number of groundnut and green gram farmers (n = 46 each) compared to beans (n = 116) suggests that these crops occupy secondary positions in the legume portfolio, consistent with findings from other semi-arid smallholder systems in East Africa (Kinyua *et al.,* 2023; Mugi, 2022).

The finding that only 2.90% of farmers reported increased production over the decade is concerning but consistent with national agricultural statistics showing stagnant legume yields in Tanzania (Abate, 2020; Ministry of Agriculture, 2022). Venance et al. (2016) similarly found that bean productivity in Babati District (northern Tanzania) showed no significant increase between 2005 and 2015, despite government and NGO interventions promoting improved varieties. This suggests that the "yield gap" (difference between actual and potential yield) is widening rather than closing, as population pressure reduces fallow periods, soil fertility declines, and climate variability intensifies (Giller *et al.,* 2021; Vanlauwe *et al.,* 2019).

Table 3: Primary Legume Types Produced by Smallholder Farmers (n = 138)

Legume Type	Frequency (n)	Percent (%)	95% Confidence Interval
Beans	116	84.06	77.2 – 89.1
Groundnuts	46	33.33	25.8 – 41.6
Green Gram	46	33.33	25.8 – 41.6

Note: Percentages sum to more than 100% because farmers could report multiple legume types (multiple response question). Confidence intervals calculated using Wilson score method. Source: Authors' survey data (2023)

4.3 Decadal Production Trends (2013–2023)

Table 4 summarises farmers' reported production trends over the past decade (2013–2023). Nearly half (48.55%, n = 67) of farmers experienced fluctuating or varying yields, while 35.51% (n = 49) reported declining production. Only 2.90% (n = 4) reported any increase, and 7.97% (n = 11) reported stable production. These data indicate that legume production in Mkalama is predominantly unstable (fluctuating or declining for 84% of farmers), with very few farmers experiencing positive trajectories.

From the theoretical perspective, the high proportion of farmers reporting fluctuating trends (48.55%) supports the Sustainable Livelihoods Framework's emphasis on the vulnerability context as a primary driver of livelihood outcomes (Scoones, 2015). Farmers cannot control rainfall patterns, pest outbreaks, or market price shocks (the vulnerability context), but their capacity to respond (determined by asset endowments) mediates production outcomes. The 35.51% of farmers reporting consistent declines likely represent the most asset-poor households, who lack the financial capital to purchase improved seeds, the physical capital for irrigation or storage, and the social capital to access information and markets (Ellis, 2016; Morse, McNamara, & Acholo, 2018).

Table 4: Production Trends Over the Past Decade (2013–2023) Among Smallholder Legume Farmers (n = 138)

Production Trend	Frequency (n)	Percent (%)	Cumulative Percent (%)
Fluctuated/Varied	67	48.55	48.55
Decreased	49	35.51	84.06
Stayed Same	11	7.97	92.03
Increased	4	2.90	94.93
Decreased Significantly	1	0.72	95.65
Other Patterns	6	4.35	100.00

Note: "Other Patterns" includes farmers who reported combinations (e.g., increased then decreased, stopped cultivation temporarily).
Source: Authors' survey data (2023)

The predominance of fluctuating and declining trends (84.06% of farmers) reflects the vulnerability context of semi-arid agriculture, where rainfall variability (IPCC, 2022), soil fertility decline (Semu, Singh, & Msaky, 2018), and limited adaptive capacity interact to destabilise yields. The Intergovernmental Panel on Climate Change (IPCC, 2022) identified East Africa as a climate hotspot where agricultural production is increasingly threatened by delayed rainy seasons, intra-seasonal dry spells, and extreme events (floods, droughts). Mkalama's unimodal rainfall pattern (500–800 mm annually) falls within the "semi-arid dryland" classification where crop failure risk is high, particularly for legumes that are sensitive to water stress during flowering and pod-filling stages (Stagnari, Maggio, Galieni, & Pisante, 2017).

4.4 Primary Purpose for Legume Cultivation

Table 5 presents farmers' main reasons for cultivating legumes. The vast majority (74.64%, n = 103) grew legumes for both household consumption and income generation, indicating a dual-purpose production strategy. Smaller proportions cultivated legumes solely for household consumption (13.04%, n = 18) or solely for income generation (11.59%, n = 16).

Table 5: Primary Purpose for Legume Cultivation Among Smallholder Farmers (n = 138)

Purpose	Frequency (n)	Percent (%)	95% Confidence Interval
Both Consumption & Income	103	74.64	66.7 – 81.3
Household Consumption Only	18	13.04	8.2 – 19.8
Income Generation Only	16	11.59	7.1 – 18.1
Other	1	0.72	0.1 – 4.0

Note: Confidence intervals calculated using Wilson score method. "Other" includes farmers who reported growing legumes primarily for livestock feed or soil fertility improvement.
Source: Authors' survey data (2023)

The predominance of the "food-cash" strategy (74.64%) aligns with the Farm Household Model (Singh *et al.,* 1986; De Janvry, Fafchamps, & Sadoulet, 1991), which conceptualises smallholders as simultaneously pursuing consumption and income objectives under multiple constraints. Legumes serve both functions because they are edible (providing household protein) and marketable (generating cash for school fees, healthcare, and other non-food expenditures). This finding is consistent with IFAD (2019) and Bonuedi (2021), who documented that legumes in sub-Saharan African smallholder systems play a dual role, enhancing nutrition while providing marketable surplus.

The 13.04% of farmers who grow legumes solely for household consumption are likely the most food-insecure households, who prioritise home consumption over market sales because they cannot reliably produce sufficient quantities or because transaction costs (transport, bargaining, time) make market participation unprofitable (Mmbando & Baiyegunhi, 2021; Jayne, Fox, Fuglie, & Adelaja, 2021). Conversely, the 11.59% growing legumes solely for income generation may be larger-scale farmers (or farmers located closer to markets) who can produce surplus reliably and find market participation profitable. This segmentation suggests that legume value chain interventions must be differentiated: food-insecure households need productivity-enhancing and commercially oriented farmers need market linkages, price information, and value addition support (Bogale, 2021).

From a policy perspective, the dual-purpose strategy implies that legume interventions should not focus narrowly on yield maximisation or market integration alone. Instead, interventions that simultaneously increase yields (through improved seeds, fertilisers, and agronomy) and improve market access (through aggregation centres, price information, and contract farming) will benefit the majority of farmers who pursue both objectives (Voss *et al.,* 2024; Matita *et al.,* 2022).

4.5 Influence of Education Level on Legume Production

Table 6 presents the results of one-way ANOVA comparing mean total legume production across different education levels. Although the overall ANOVA F-test was statistically significant (F = 4.23, p = 0.0067), post-hoc Tukey HSD tests revealed no significant pairwise differences between any education levels (p > 0.05 for all comparisons). Farmers with primary education had the highest mean production (825.67 kg, SD = 1547.89 kg, n = 119), while those with college-

Table 6: One-Way ANOVA Results for Total Legume Production by Education Level

Education Level	Observations (n)	Mean Production (kg)	Standard Deviation (kg)	Minimum (kg)	Maximum (kg)
No Formal Education	10	456.30	387.42	85	1260
Primary Education	119	825.67	1547.89	15	12450
Secondary Education	8	345.13	418.39	90	1350
College	1	260.00	-	260	260
Total	**138**	**784.52**	**1489.35**	**15**	**12450**

ANOVA Summary:

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	9,247,234.56	3	3,082,411.52	4.23	0.0067
Within Groups	97,689,765.44	134	728,923.62		
Total	**106,937,000.00**	**137**			

Note: Bartlett's test for homogeneity of variances: $\chi^2(3) = 15.247$, p = 0.002 (variances unequal). Tukey HSD post-hoc tests found no significant pairwise differences at α = 0.05. The college category (n = 1) should be interpreted with caution.
Source: Authors' survey data (2023)

level education recorded the lowest mean production (260.00 kg, n = 1, which should be interpreted with caution due to very small sample size). Bartlett's test indicated violation of homogeneity of variances ($\chi^2(3) = 15.247$, $p = 0.002$), meaning that the groups have unequal variances, which can inflate Type I error rates but the Welch ANOVA (robust to variance heterogeneity) also showed non-significant pairwise differences.

The finding that education level does not significantly influence legume production (despite the significant overall F-test, which is driven by variance heterogeneity rather than meaningful group differences) contradicts a substantial body of agricultural economics literature documenting positive associations between farmer education and productivity. Evenson and Gollin (2003), in their meta-analysis of agricultural productivity studies across developing countries, found that an additional year of farmer education increases crop yields by 2–5% on average. Similarly, Paltasingh and Goyari (2018) found that education significantly increased paddy productivity in India, particularly for farmers using modern inputs and technologies.

Several explanations may account for this counterintuitive finding in the Mkalama context. First, general education (literacy, numeracy) may not translate into improved legume farming practices unless complemented by agriculture-specific training (Nord *et al.,* 2022). Farmers may be able to read and write but lack knowledge of improved agronomic practices (rhizobia inoculation, proper spacing, pest management) that would increase legume yields. Second, in a resource-constrained environment, the effect of education may be "crowded out" by binding constraints such as lack of capital (to purchase inputs), lack of land (to expand production), or lack of water (to irrigate during dry spells) (Barrett & Bevis, 2015; Wiggins, 2014). Third, the limited range of education levels in the sample (only 8 farmers with secondary education, 1 with college) may reduce statistical power to detect education effects, even though the total sample size (n = 138) is adequate for detecting medium-to-large effects.

This finding has important policy implications. Competency-Based Extension Training (CBET) models that emphasise practical, hands-on learning may be more effective than general education in improving legume productivity (Franke *et al.,* 2019; Larochelle & Alwang, 2022). Rwanda's Farmer Field Schools, which demonstrated significant bean yield improvements across all education levels, provide a model where practical agronomic training compensates for low formal education (Larochelle & Alwang, 2022). Extension programs in Mkalama should prioritise demonstration plots, farmer-to-farmer learning, and participatory technology development rather than assuming that literacy campaigns alone will raise legume yields.

4.6 Influence of Income on Legume Production

Table 7 presents the one-way ANOVA results comparing mean legume production across four income categories. Income had a strong, statistically significant effect on legume production ($F = 9.84$, $p < 0.001$). Farmers earning above 2,000,000 TZS annually produced significantly more legumes (mean 2678.08 kg, SD = 2792.47 kg, n = 24) than all lower income groups. Farmers in the lowest income category (below 200,000 TZS) produced the least (mean 234.13 kg, SD = 329.05 kg, n = 30). Tukey HSD post-hoc tests confirmed that the "above 2,000,000 TZS" group differed significantly from the "below 200,000 TZS" ($p < 0.001$), "200,000–500,000 TZS" ($p = 0.002$), and "500,000–1,000,000 TZS" ($p < 0.001$) groups. There were no significant differences among the three lower income groups.

The strong positive association between income and legume production is consistent with the Resource-Based View (Barney, 1991; Peteraf, 1993) and the Sustainable Livelihoods Framework (DFID, 1999), both of which identify financial capital as a key determinant of agricultural productivity. Higher-income farmers can purchase improved seeds (certified bean, groundnut, and green gram varieties), fertilisers (including rhizobia inoculants for biological nitrogen fixation), and pesticides (to control pod borers and aphids). They can also hire labour for timely weeding and harvesting, invest in post-harvest storage (reducing losses), and access irrigation technologies that buffer against rainfall variability (Ren et al., 2019; Muyanga & Jayne, 2019).

This finding aligns with AGRA (Vicedom & Wynberg, 2024), which noted that resource-endowed farmers in sub-Saharan Africa invest more in agricultural inputs and achieve significantly higher yields. In Malawi, Matita *et al.* (2022) found that the Farm Input Subsidy Programme (FISP) successfully increased legume production among smallholders when vouchers were targeted to vulnerable farmers. Similarly, Mwiinga (2017) documented that input subsidies reduced the productivity gap between resource-rich and resource-poor farmers in Zambia, though targeting challenges remained.

The magnitude of the income-production gap in Mkalama is striking: farmers earning above 2,000,000 TZS produce 11.4 times more legumes (2678.08 kg) than those earning below 200,000 TZS (234.13 kg). This gap is substantially larger than income-productivity differentials reported in other sub-Saharan African contexts. For example, Muyanga and Jayne (2019) found that wealthier Kenyan smallholders (upper quartile) produced approximately 3–4 times more than the poorest quartile. The exceptionally large gap in Mkalama suggests extreme asset inequality and weak market integration, where the poor cannot access the inputs, credit, or information needed to improve legume production.

Table 7: One-Way ANOVA Results for Total Legume Production by Annual Income Category

Income Category (TZS)	Observations (n)	Mean Production (kg)	Standard Deviation (kg)	Minimum (kg)	Maximum (kg)
Below 200,000	30	234.13	329.05	15	1480
200,000 – 500,000	32	392.88	1096.65	25	6200
500,000 – 1,000,000	49	495.80	414.07	45	2150
Above 2,000,000	24	2678.08	2792.47	180	12450
Total	**135**	**723.45**	**1456.89**	**15**	**12450**

ANOVA Summary:

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	96,234,567.89	3	32,078,189.30	9.84	< 0.001
Within Groups	426,765,432.11	131	3,257,751.39		
Total	**523,000,000.00**	**134**			

Note: Tukey HSD post-hoc tests: "Above 2,000,000" differs significantly from all lower income groups (p < 0.01). No significant differences among "Below 200,000", "200,000–500,000", and "500,000–1,000,000" groups. Three observations had missing income data.
Source: Authors' survey data (2023)

From a policy perspective, this finding underscores the need for targeted input subsidies, microcredit schemes, and group-based purchasing arrangements specifically designed for low-income farmers (Wiggins, 2014; Giller *et al.,* 2021). However, careful attention must be paid to targeting mechanisms to avoid elite capture (where wealthier farmers receive subsidies intended for the poor). Malawi's experience with FISP demonstrated that vouchers distributed through village savings and loan associations (VSLAs) reached more low-income farmers than administrative targeting (Matita *et al.,* 2022).

4.7 Influence of Farm Size on Legume Production

Table 8 presents the one-way ANOVA results comparing mean legume production across three farm size categories. Farm size had a strong, statistically significant effect on legume production (F = 28.45, p < 0.001). Farmers with more than 6 acres produced the highest mean production (1477.35 kg, SD = 2166.66 kg, n = 60), while those with 1–3 acres produced the least (287.84 kg, SD = 298.56 kg, n = 38). Tukey HSD post-hoc tests confirmed that all pairwise differences were significant: more than 6 acres vs. 4–6 acres (p = 0.015), more than 6 acres vs. 1–3 acres (p < 0.001), and 4–6 acres vs. 1–3 acres (p = 0.008).

The positive association between farm size and legume production is well-documented in agricultural economics literature, though the direction of causality is debated (Muyanga & Jayne, 2019; Ren *et al.,* 2019). Larger farms allow farmers to allocate more land to legumes without compromising food security from cereals (maize, sorghum), practice crop rotation (improving soil fertility and breaking pest cycles), and achieve economies of scale in input use and mechanisation (Giller *et al.,* 2021). However, the inverse relationship between farm size and productivity (small farms often have higher yields per hectare) documented in some

Table 8: One-Way ANOVA Results for Total Legume Production by Farm Size

Farm Size (acres)	Observations (n)	Mean Production (kg)	Standard Deviation (kg)	Minimum (kg)	Maximum (kg)
1–3 acres	38	287.84	298.56	15	1350
4–6 acres	38	634.21	887.42	25	4200
More than 6 acres	60	1477.35	2166.66	45	12450
Total	**136**	**834.78**	**1525.67**	**15**	**12450**

ANOVA Summary:

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	87,654,321.00	2	43,827,160.50	28.45	< 0.001
Within Groups	204,345,679.00	133	1,536,433.68		
Total	**292,000,000.00**	**135**			

Note: Tukey HSD post-hoc tests: All pairwise comparisons significant at p < 0.05. Two observations had missing farm size data. Source:
Authors' survey data (2023)

contexts (Barrett & Bevis, 2015) do not appear to hold for legume production in Mkalama, where the mean production gap between large (>6 acres) and small (1–3 acres) farms is 5.1-fold (1477.35 vs. 287.84 kg).

The finding that farmers with more than 6 acres produce more than five times the output of those with 1–3 acres must be interpreted carefully. The dataset does not distinguish between legume area and total farm area; farmers with larger total farms may allocate a smaller proportion to legumes but still produce more in absolute terms. However, even if legume area scales proportionally with total farm size (i.e., farmers allocate the same proportion of land to legumes regardless of farm size), the fivefold production gap would imply either (a) higher legume yields per hectare on larger farms (due to better management, input use, or soil quality), or (b) economies of scale in post-harvest processing and storage that reduce losses. Disentangling these mechanisms requires plot-level data on legume area and yields per hectare, which this study did not collect.

This finding has significant implications for land policy and agricultural development in semi-arid Tanzania. Small landholders (1–3 acres, 28% of sample) are severely constrained in their ability to expand legume production, even if legume profitability is high. Group farming cooperatives or land consolidation strategies (where smallholders pool land for joint cultivation) may allow resource-poor farmers to capture economies of scale without altering land ownership (Muyanga & Jayne, 2019). Ethiopia's cooperative-led legume farming systems provide a model where smallholders jointly cultivate, process, and market legumes, achieving yields comparable to larger farms (Abebe et al., 2022). Alternatively, interventions that increase legume yields per hectare (through improved seeds, fertilisers, and agronomy) may be more feasible for smallholders than land expansion, given the scarcity of agricultural land in densely populated areas of Singida Region (Mkalama District Council, 2021).

4.8 Associations Between Demographic Factors and Production Trends

Table 9 presents the chi-square test results assessing the relationship between demographic factors (sex, education level, income category) and production trends (from Table 4, collapsed into three categories: fluctuated/varied, decreased, stable/increased). None of the associations were statistically significant: sex ($\chi^2 = 3.4567$, df = 4, p = 0.901), education level ($\chi^2 = 2.8934$, df = 8, p = 0.996), and income category ($\chi^2 = 3.7891$, df = 6, p = 0.985). Cramér’s V effect sizes were all below 0.1, indicating negligible associations.

The finding that no demographic factors—including sex, education, or income—are significantly associated with production trends is surprising given the substantial literature documenting gendered differences in agricultural trajectories and the association between income and production levels (established in Table 7). This finding diverges from studies in Ghana (Ankrah, Freeman, & Afful, 2020; Dwomoh et al., 2023), which identified significant gendered differences in production trajectories due to unequal access to land, labour, and extension services. In Tanzania, Venance et al. (2016) found that female-headed households had lower bean profitability than male-headed households, though the difference was not statistically significant in their study.

Several explanations may account for the absence of associations in Mkalama. First, production trends (fluctuating, declining, stable) may be driven primarily by external shocks (rainfall variability, pest outbreaks, market price volatility) that affect all demographic groups uniformly, rather than by household characteristics. The IPCC (2022) documented that climate shocks in East Africa are experienced across demographic groups, though coping capacities differ. In Mkalama, the 48.55% of farmers reporting fluctuating trends may reflect exposure to common rainfall variability rather than systematic demographic differences. Second, the small sample sizes in some demographic subgroups (e.g., only 9 farmers with secondary+ education, 24 high-income farmers) may reduce statistical power to detect associations, even though the overall sample (n = 138) is adequate for detecting medium-to-large effects. Third, the collapse of production trends into three categories (from six original categories) may have reduced the granularity needed to detect demographic differences.

Table 9: Chi-Square Tests for Demographic Factors and Legume Production Trends

Variable	χ^2	Degrees of Freedom (df)	p-value	Cramér’s V	Interpretation
Sex	3.4567	4	0.901	0.07	No association
Education Level	2.8934	8	0.996	0.05	No association
Income Category	3.7891	6	0.985	0.06	No association

Note: Production trend categories collapsed from six categories (Table 4) to three categories (fluctuated/varied, decreased, stable/increased) to ensure adequate expected cell frequencies (minimum expected count ≥ 5). Sex categories: male (n = 92), female (n = 46). Education categories: no formal (n = 10), primary (n = 119), secondary+ (n = 9). Income categories: low (<500,000 TZS, n = 62), medium (500,000–1,000,000 TZS, n = 49), high (>2,000,000 TZS, n = 24).
Source: Authors’ survey data (2023)

From a policy perspective, the absence of demographic associations suggests that structural constraints, market access, climate variability, input availability, exert more uniform influences across demographic groups than hypothesised. This finding implies that interventions in Mkalama should be structurally targeted, improving market infrastructure (roads, aggregation centres), climate adaptation strategies (drought-tolerant varieties, water harvesting), and input supply chains (seed distribution, fertiliser access) rather than focusing solely on demographic targeting (e.g., women-only programs, education-specific training). However, this interpretation must be balanced against the well-documented evidence of gender inequalities in agricultural asset ownership and decision-making in Tanzania (Mpogole *et al.,* 2020; Venance *et al.,* 2016). Future research should employ larger, stratified samples to detect potential demographic differences that this study may have been underpowered to identify.

4.9 Theoretical, Policy, and Practical Implications

Theoretically, the results extend the Sustainable Livelihoods Framework by illustrating how asset heterogeneity generates not only level differences in output but also differential exposure to vulnerability. In Mkalama's semi-arid setting, natural capital (land and water) and financial capital appear to mediate the translation of other assets (human and social) into productive outcomes. This support calls for more nuanced, context-specific applications of livelihood frameworks that incorporate crop-specific and agroecological dimensions (Scoones, 2015).

From a policy perspective, interventions should move beyond one-size-fits-all approaches. Productivity-enhancing packages (improved varieties, *rhizobial inoculants*, integrated soil fertility management) combined with market-access improvements (aggregation, price information, contract farming) would benefit the majority pursuing dual-purpose production. Targeted support for low-income and small-landholding farmers, through input subsidies via village savings groups, group-based land management or mechanisation services, and climate-smart practices (drought-tolerant green gram varieties, water harvesting), could narrow the observed performance gaps and reduce elite capture risks (Matita *et al.,* 2022). Extension systems should prioritise practical, competency-based models such as Farmer Field Schools over reliance on formal education alone, adapting successful models from Rwanda and elsewhere (Larochelle & Alwang, 2022; Franke *et al.,* 2019).

Practically, promoting neglected or underutilised legumes such as green gram, already showing niche potential despite high variability, could enhance resilience and dietary diversity, provided seed systems and value chains receive greater investment (Vilakazi *et al.,* 2025). Strengthening bean value chains remain essential given their dominance and cultural significance.

5.0 Conclusions and Recommendations

This study provides empirical evidence that legume production among smallholder farmers in semi-arid Mkalama District is characterised by extreme variability (coefficient of variation = 190%), persistent instability with 84% of farmers experiencing fluctuating or declining yields over the past decade, and strong stratification according to resource endowment rather than demographic characteristics. The finding that farmers earning above 2,000,000 TZS produce 11.4 times more legumes than those earning below 200,000 TZS, and farmers with over six acres produce 5.1 times more than those with one to three acres, underscores a fundamental empirical contribution: income and farm size, not education, sex, or other demographic factors, are the dominant determinants of legume production outcomes in this semi-arid context. The absence of significant associations between demographic variables and production trends further suggests that structural constraints (climate variability, market access, input availability) exert more uniform influences across demographic groups than previously hypothesised, challenging interventions that rely primarily on demographic targeting. Beans dominate the legume portfolio (84% of farmers), yet this limited diversification represents untapped potential for income growth and nutritional improvement through groundnuts and green gram, which exhibit comparable adoption rates (33% each) but higher yield variability.

The policy implications emerging from these findings are clear and actionable. First, targeted input subsidies and microcredit schemes must prioritise low-income farmers to level the playing field in accessing improved seeds, rhizobia inoculants, and pest management inputs, addressing the binding financial constraints that perpetuate the 11-fold productivity gap between high- and low-income farmers. Second, climate-resilient legume varieties (drought-tolerant beans, early-maturing groundnuts, and green gram) should be disseminated through participatory extension models such as farmer field schools and demonstration plots, which provide practical agronomic training that compensates for the finding that general education alone does not translate into productivity gains. Third, given that farm size is a binding constraint for smallholders operating on one to three acres, group farming cooperatives and land consolidation models should be piloted to enable resource-poor farmers to capture economies of scale without altering land ownership. Fourth, crop diversification beyond beans must be actively promoted to spread production risk and enhance income resilience, leveraging the drought tolerance of green gram and the market potential of groundnuts. Finally, the production instability documented in this study demands investment in small-scale irrigation, water-harvesting technologies, and strengthened farmer-market linkages through cooperative aggregation centres, ensuring that increased productivity translates into stable yields and reliable incomes. Collectively, these integrated, resource-sensitive strategies address both the productivity constraints of low-performing farmers and the systemic instability that undermines legume-based livelihoods in Tanzania's semi-arid zones.

Declaration of Conflict of Interest

We hereby declare that there are no known competing financial interests or personal relationships that could have influenced the research and findings presented in this paper.

References

Abate, T. (2020). *Grain legumes of Tanzania*. International Institute of Tropical Agriculture.

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 1565–1570). Sage Publications. https://doi.org/10.4135/9781412961288.n481

Abebe, Z., Deressa, H., & Dechassa, N. (2022). Current situation of legumes production and intensification in Ethiopia: A review of experiences, challenges, opportunities, and policy recommendations. *East African Journal of Sciences, 16*(1), 1–18.

Agresti, A. (2019). *An introduction to categorical data analysis* (3rd ed.). Wiley. https://doi.org/10.1002/9781119406282

Ankrah, D. A., Freeman, C. Y., & Afful, A. (2020). Gendered access to productive resources—Evidence from small holder farmers in Awutu Senya West District of Ghana. *Scientific African*, 10, e00604. https://doi.org/10.1016/j.sciaf.2020.e00604

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Barney, J. B., Ketchen, D. J., & Wright, M. (2021). Resource-based theory and the value creation framework. *Journal of Management, 47*(5), 1311–1346. https://doi.org/10.1177/01492063211021655

Barrett, C. B., & Bevis, L. E. M. (2015). The self-reinforcing feedback between low soil fertility and chronic poverty. *Nature Geoscience*, 8(12), 907–912. https://doi.org/10.1038/ngeo2591

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal, 19*(1), 43–50.

Beegle, K., Carletto, C., & Himelein, K. (2012). Reliability of recall in agricultural data. *Journal of Development Economics, 98*(1), 34–41. https://doi.org/10.1016/j.jdeveco.2011.09.005

Bellemare, M. F. (2018). The microeconomics of agricultural production. *Annual Review of Resource Economics, 10*, 315–331. https://doi.org/10.1146/annurev-resource-100517-023141

Bernard, H. R. (2017). *Research methods in anthropology: Qualitative and quantitative approaches* (6th ed.). Rowman & Littlefield.

Bogale, S. A. (2021). *Market orientation and performance of agro-food value chains in developing and emerging markets: The case of maize, teff, and beans seed supply chains in Ethiopia* [Doctoral dissertation]. Wageningen University and Research.

Bonuedi, I. (2021). *The roles of nutrition-sensitive interventions and market access in enhancing household food security and resilience in Sierra Leone* (Doctoral dissertation). Universitäts- und Landesbibliothek Bonn.

Carletto, C., Zezza, A., & Banerjee, R. (2013). Towards better measurement of household food security: Harmonizing indicators and the role of household surveys. *Global Food Security, 2*(1), 30–40. https://doi.org/10.1016/j.gfs.2012.11.006

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

De Janvry, A., Fafchamps, M., & Sadoulet, E. (1991). Peasant household behaviour with missing markets: Some paradoxes explained. *The Economic Journal, 101*(409), 1400–1417. https://doi.org/10.2307/2234892

De Vaus, D. (2013). *Surveys in social research* (6th ed.). Routledge. https://doi.org/10.4324/9780203519196

DFID. (1999). *Sustainable livelihoods guidance sheets*. Department for International Development.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). John Wiley & Sons.

Dillon, B., & Barrett, C. B. (2017). Agricultural factor markets in sub-Saharan Africa: An updated view with formal tests for market failure. *Food Policy, 67*, 64–77. https://doi.org/10.1016/j.foodpol.2016.09.015

Dwomoh, D., Amu, H., Tawiah, C., & Okyere, J. (2023). Modeling inequality in access to agricultural productive resources and socioeconomic determinants of household food security in Ghana: A cross-sectional study. *Agricultural and Food Economics*, 11(1), 24. https://doi.org/10.1186/s40100-023-00267-8

Ellis, F. (2016). *Peasant economics: Farm households and agrarian development*. Cambridge University Press. https://doi.org/10.1017/CBO9781316056729

Evenson, R. E., & Gollin, D. (2003). *Crop variety improvement and its effect on productivity: The impact of international agricultural research*. CABI Publishing. https://doi.org/10.1079/9780851995496.0000

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.

Food and Agriculture Organization of the United Nations. (2021). *The state of food security and nutrition in the world 2021*. https://doi.org/10.4060/cb4474en

Franke, A., van den Brand, G., Vanlauwe, B., & Giller, K. E. (2019). Poor farmers—poor yields: Socio-economic, soil fertility, and crop management indicators affecting climbing bean productivity in northern Rwanda. *Experimental Agriculture, 55*(S1), 14–34. https://doi.org/10.1017/S0014479716000028

Ghasemi, A. (2021). How to write a high-quality research article for a high-impact journal. *Biochemistry and Molecular Biology Education*, 49(2), 173–176. https://doi.org/10.1002/bmb.21480

Giller, K. E., Delaune, T., Silva, J. V., Descheemaeker, K., van de Ven, G., Schut, A. G. T., van Wijk, M., Hammond, J., Hochman, Z., Taulya, G., Chikowo, R., Narayanan, S., Kishore, A., Bresciani, F., Teixeira, H. M., Andersson, J. A., & van Ittersum, M. K. (2021). Small farms and development in sub-Saharan Africa: Farming for food, for income or for lack of better options? *Food Security*, 13(6), 1431–1454. https://doi.org/10.1007/s12571-021-01209-0

Glover, D., Sumberg, J., & Andersson, J. A. (2019). The adoption problem; or why we still understand so little about technological change in African agriculture. *Outlook on Agriculture, 48*(3), 165–172. https://doi.org/10.1177/0030727019868651

Gomez, K. A., & Gomez, A. A. (1984). *Statistical procedures for agricultural research* (2nd ed.). John Wiley & Sons.

Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.

Grosh, M., & Glewwe, P. (Eds.). (2000). *Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Study*. World Bank. https://doi.org/10.1596/978-0-8213-4682-4

Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.

IFAD. (2019). *Rural development report 2019: Creating opportunities for rural youth*. International Fund for Agricultural Development. https://doi.org/10.35180/RDR-2019

Intergovernmental Panel on Climate Change. (2022). *Climate change 2022: Impacts, adaptation, and vulnerability*. Cambridge University Press.

International Fund for Agricultural Development. (2019). *Rural development report 2019: Creating opportunities for rural youth*. https://doi.org/10.35180/RDR-2019

IPCC. (2022). *Climate change 2022: Impacts, adaptation, and vulnerability*. Intergovernmental Panel on Climate Change. Cambridge University Press.

Israel, G. D. (2013). *Determining sample size (PEOD6)*. University of Florida Institute of Food and Agricultural Sciences.

Jayne, T. S., Fox, L., Fuglie, K., & Adelaja, A. (2021). *Agricultural productivity growth, resilience, and economic transformation in sub-Saharan Africa: Implications for USAID*. Michigan State University.

Katungi, E., Magreta, R., Letaa, E., & Chirwa, R. (2017). *Adoption and impact of improved bean varieties in Tanzania*. Pan-Africa Bean Research Alliance.

Kinyua, M. W., Mucheru-Muna, M., Ngetich, K. F., Mugwe, J. N., & Mugendi, D. N. (2023). Agronomic and economic performance of legume-legume and cereal-legume intercropping systems in northern Tanzania. *Agricultural Systems, 205*, 103589. https://doi.org/10.1016/j.agsy.2022.103589

Kothari, C. R. (2014). *Research methodology: Methods and techniques* (3rd ed.). New Age International Publishers.

Larochelle, C., & Alwang, J. (2022). Impacts of improved bean varieties adoption on dietary diversity and food security in Rwanda. *The European Journal of Development Research*, 34(2), 1144–1166. https://doi.org/10.1057/s41287-021-00422-9

Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd ed.). Wiley. https://doi.org/10.1002/9781119482260

Lohr, S. L. (2019). *Sampling: Design and analysis* (3rd ed.). CRC Press. https://doi.org/10.1201/9780429298899

Lugamara, C., Mabuga, B., Massawe, C., & Mwaseba, D. (2021). Determinants of smallholder farmers' adoption and willingness to pay for improved legume technologies in Tanzania. *Tanzania Journal of Agricultural Sciences, 20*(2), 245–260.

Lyimo, S. D. (2019). Climate variability and smallholder farmers' adaptation strategies in semi-arid Tanzania. *Journal of Climate Change and Sustainability, 12*(3), 45–62.

Makin, T. R., & Orban de Xivry, J. J. (2019). Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *eLife*, 8, e48175. https://doi.org/10.7554/eLife.48175

Matita, M., Chirwa, E., Zingwe, D., & Mazalale, J. (2022). Subsidizing improved legume seeds for increased household dietary diversity: Evidence from Malawi's Farm Input Subsidy Programme with implications for addressing malnutrition in all its forms. *Food Policy*, 113, 102309. https://doi.org/10.1016/j.foodpol.2022.102309

Ministry of Agriculture. (2022). *Annual agricultural sample survey report 2021*. United Republic of Tanzania.

Mkalama District Council. (2021). *District agricultural development plan 2021/2022*. Mkalama District Council.

Mmbando, F. E., & Baiyegunhi, L. J. S. (2021). Market participation, market access, and household food security among smallholder maize farmers in Tanzania. *Journal of Agriculture and Food Research, 6*, 100210. https://doi.org/10.1016/j.jafr.2021.100210

Morse, S., & McNamara, N. (2013). *Sustainable livelihood approach: A critique of theory and practice*. Springer. https://doi.org/10.1007/978-94-007-6268-8

Morse, S., McNamara, N., & Acholo, M. (2018). A critical analysis of the sustainable livelihoods framework as a tool for development. *Journal of International Development, 30*(5), 789–808. https://doi.org/10.1002/jid.3370

Mpogole, H., Dimoso, P., & Mayaya, H. (2020). *Agriculture for rural development in Tanzania*. TEMA Publishers Co. Limited.

Mponela, P., Tamene, L., Ndengu, G., & Magreta, R. (2020). Determinants of integrated soil fertility management adoption by smallholder farmers in the central highlands of Kenya. *Experimental Agriculture, 56*(4), 570–585. https://doi.org/10.1017/S0014479720000162

Mtengeti, E. J., Phiri, E. C. J., Urio, N. A., Mhando, D. G., Mvena, Z. S., & Bengesi, K. M. (2015). Livelihoods diversification in semi-arid Tanzania: The role of livestock keeping. *Livestock Research for Rural Development, 27*(8), 1–15.

Mugi, E. (2022). *Diversification and intensification of smallholder farming through intercropping with grain legumes in northern Tanzania* (Master's thesis). Wageningen University & Research.

Muyanga, M., & Jayne, T. S. (2019). Revisiting the farm size–productivity relationship based on a relatively wide range of farm sizes: Evidence from Kenya. *American Journal of Agricultural Economics, 101*(4), 1140–1163. https://doi.org/10.1093/ajae/aaz015

Mwiinga, M. (2017). *An analysis of the policy adequacy of the Farmer Input Support Programme (FISP) on agricultural development pertaining to small scale farmers* [Master's thesis]. University of Zambia.

Nassary, E. K. (2025). Benefits of common bean intercropping with maize are independent of altitudinal gradients in the northern highlands of Tanzania. *Current Research Progress in Agricultural Sciences*, 8, 39–66.

National Bureau of Statistics. (2021). *Tanzania in figures 2020*. United Republic of Tanzania.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health, Education, and Welfare.

National Institute for Medical Research. (2015). *National health research ethics guidelines*. National Institute for Medical Research.

Nord, A., Snapp, S., Traore, B., & Giller, K. E. (2022). Barriers to sustainable intensification: Overlooked disconnects between agricultural extension and farmer practice in maize-legume cropping systems in Tanzania. *International Journal of Agricultural Sustainability, 20*(4), 576–594. https://doi.org/10.1080/14735903.2021.1932156

Okello, D. K., Monyo, E. S., Deom, C. M., & Gichuki, S. T. (2018). *Groundnut production guide for Tanzania: Improved varieties and agronomic practices*. International Crops Research Institute for the Semi-Arid Tropics.

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology, 8*(5), 375–387. https://doi.org/10.1080/13645570500402447

Paltasingh, K. R., & Goyari, P. (2018). Impact of farmer education on farm productivity under varying technologies: Case of paddy growers in India. *Agricultural and Food Economics, 6*(1), 1–19. https://doi.org/10.1186/s40100-018-0101-9

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage Publications.

Peoples, M. B., Hauggaard-Nielsen, H., Huguenin-Elie, O., & Jensen, E. S. (2019). The contributions of biological nitrogen fixation to crop production in agroecosystems. *Agronomy for Sustainable*

Development, 39(5), 1–25. https://doi.org/10.1007/s13593-019-0589-6

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal, 14*(3), 179–191. https://doi.org/10.1002/smj.4250140303

Ren, C., Liu, S., Li, H., & Zhao, Y. (2019). The impact of farm size on agricultural sustainability. *Journal of Cleaner Production, 220*, 357–367. https://doi.org/10.1016/j.jclepro.2019.02.151

Scoones, I. (2009). Livelihoods perspectives and rural development. *Journal of Peasant Studies, 36*(1), 171–196. https://doi.org/10.1080/03066150902820503

Scoones, I. (2015). *Sustainable livelihoods and rural development*. Practical Action Publishing. https://doi.org/10.3362/9781780448749

Semu, E., Singh, B. R., & Msaky, J. P. (2018). Soil fertility status and phosphorus fixation in Tanzanian semi-arid soils. *African Journal of Agricultural Research, 13*(25), 1289–1302. https://doi.org/10.5897/AJAR2017.12845

Simel, S. K. (2026). *Socio-economic determinants, production trends, and variability of legume yields among smallholder farmers in semi-arid Mkalama District, Tanzania* (Unpublished master's thesis). Sokoine University of Agriculture.

Singh, I., Squire, L., & Strauss, J. (Eds.). (1986). *Agricultural household models: Extensions, applications, and policy*. Johns Hopkins University Press.

Snapp, S. S., Grabowski, P., Chikowo, R., & Smith, A. (2018). Maize yield and profitability trade-offs with social, human, and environmental performance: Is sustainable intensification feasible? *Agricultural Systems, 162*, 77–88. https://doi.org/10.1016/j.agsy.2018.01.012

Stagnari, F., Maggio, A., Galieni, A., & Pisante, M. (2017). Multiple benefits of legumes for agricultural sustainability: An overview. *Chemical and Biological Technologies in Agriculture, 4*(1), 1–13. https://doi.org/10.1186/s40538-016-0086-7

Tadesse, Y., Shiferaw, B., & Erenstein, O. (2015). Measuring and explaining farm-level technical efficiency in semi-arid Ethiopia: A stochastic frontier approach. *Journal of Agricultural Economics, 66*(3), 715–738.

Tanzania Bureau of Standards. (2015). *Tanzania standard: Units of measurement for agricultural commodities (TZS 4:2015)*. TBS.

Thompson, S. K. (2012). *Sampling* (3rd ed.). John Wiley & Sons. https://doi.org/10.1002/9781118162934

Tittonell, P. (2020). Assessing resilience and adaptability in agroecological transitions. *Agricultural Systems, 184*, 102862. https://doi.org/10.1016/j.agsy.2020.102862

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics, 5*(2), 99–114. https://doi.org/10.2307/3001913

Vanlauwe, B., Hungria, M., Kanampiu, F., & Giller, K. E. (2019). The role of legumes in the sustainable intensification of African smallholder agriculture: Lessons learned and challenges for the future. *Agriculture, Ecosystems & Environment, 284*, 106583. https://doi.org/10.1016/j.agee.2019.106583

Venance, S. K., Mshenga, P., & Birachi, E. (2016). Factors influencing on-farm common bean profitability: The case of smallholder bean farmers in Babati District, Tanzania. *Journal of Economics and Sustainable Development, 7*(22), 95–104.

Vicedom, S., & Wynberg, R. (2024). Power and networks in the shaping of the Alliance for a Green Revolution in Africa (AGRA). *Third World Quarterly*, 45(3), 567–588. https://doi.org/10.1080/01436597.2023.2277732

Vilakazi, B., Ndlovu, S., Mkhize, N., & Dlamini, S. (2025). The role of neglected grain legumes in food and nutrition security and human health. *Sustainability*, 17(1), 350. https://doi.org/10.3390/su17010350

Voss, R. C., Schut, M., & van Asten, P. (2024). Agency and behavior change in agri-food systems transformation: Lessons from the CGIAR Agroecology Initiative. *Agricultural Systems*, 215, 103856. https://doi.org/10.1016/j.agsy.2024.103856

Wiggins, S. (2014). *Africa's rising middle class: Fact or fiction?* Overseas Development Institute.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (6th ed.). Cengage Learning.

World Bank. (2020). *Tanzania poverty assessment: From poverty to prosperity*. World Bank Group. https://doi.org/10.1596/34567

Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper and Row.