Assessing Gendered Climate Related Vulnerabilities in Community Based Conservation Areas of Southern Sub-Saharan Africa

Andrea Clement
Department of Geography, University of Colorado Boulder

Defended April 10, 2020

Committee Members:
Sarah Kelly, Geography (Thesis Advisor)
Joel Hartter, Environmental Studies Program
Bill Travis, Honors Council Representative, Geography

Table of Contents

List of Maps, Tables, Figures, Equations, and Appendices

Page #

Maps

Tables

Figures

Equations

Appendices

Abstract

In Sub-Saharan Africa, rural households are exposed to extreme climate events, which threaten the natural resources and agricultural yields that help support livelihoods. Literature supports that vulnerability is compounded in rural female-headed households, which demonstrate less indicators of adaptive capacity when compared to male-headed counterparts. The presented research investigates if gender-based differences in vulnerability persist in community-based conservation areas (CBC's), which seek to support human livelihoods while preserving environmental and natural resources. The Kavango Zambezi Vulnerability and Adaption Project (KAZAVA) team conducted 720 household livelihood surveys in Zambia, Botswana, and Namibia to evaluate information related to income, material wealth, agriculture, natural resource use, and socio-demographic information. Using this data, I have compiled a Livelihood Vulnerability Index (LVI) to encapsulate the many dimensions of climate vulnerability. Significant differences between female and male headed households were assessed by applying a Welch's t-test, analysis of variance, and Tukey honestly significant difference (HSD) tests to LVI scores. The results of this study indicate that while some expected disparities persist within the study area, others do not. I explore these results in relationship to the potential role of CBC's in reducing disparities through generation of additional employment. This study aims to deepen the understanding of the relationship between gendered vulnerabilities and climate, to identify vulnerable populations within our studied communities, and to provide insight on how policy and employment opportunity can be used to promote climate justice in rural areas.

Research Questions:

A. How do regional patterns of gendered disparities in climate-related vulnerabilities show up in households located in community-based conservation areas (CBC's)?

B. To what extent might employment by CBC's buffer expected disparities?

1. Introduction

In the spirit of disaster triage, it is crucial to first identify who will be the most impacted by climate change. Considered a "defining issue of our time" in statements issued by the United Nations (UN), climate change continues to gain amassed attention globally (United Nations, 2016). It is widely recognized by the international community that increased atmospheric CO^2 levels and related increases in drought, flooding, and fires are of human origin (USDC, 2020; Pachauri et al., 2015). Globally, specific communities and households demonstrate less adaptive capacity, or ability to recover from and sustain wellbeing after disturbances (Kumssa & Jones, 2010; Turner et al., 2003). Relatedly, household climate vulnerability, or degree to which a household is likely to experience harm due to a climate-change related event, is decreased through adaptive capacity that may buffer complete destruction of material wealth, loss of income, and degradation of human health (Ncube et al., 2016; Turner et al., 2003).

Rural regions of Sub-Saharan Africa (SSA) are vastly composed of communities with high climate vulnerability attributed to natural resource dependency, water scarcity, and poverty (Alagidede et al., 2016; Kumssa & Jones, 2010). As SSA is not a homogenous entity, variation exists regionally, societally, and demographically, which lends specific groups to experience climatic shifts more dramatically than others (Connolly-Boutin & Smit, 2016). One such area of interest is gender disparities within rural communities. Literature supports that rural female headed households (FHH) are exceptionally susceptible to the shocks of climate change when compared to male headed households (MHH) (Alhassan et al., 2019; Ibrahim et al., 2019; Balehey et al., 2018; Assan et al., 2018; Connolly-Boutin & Smit, 2016). Studies conducted in rural areas of SSA have characterized disproportionate climate vulnerability by significant differences found between MHH and FHH in diversification of livelihood strategies, income, access to credit and financial assistance, natural resource reliance, and land and livestock ownership, among other variables (Balehey et al., 2018; Tibesigwa & Visser, 2016; Assan et al., 2018).

As recognition of vulnerable populations grows, so does the cultivation of empowering solutions. Solutions to increase adaptive capacity directed at rural communities in SSA include facilitating agricultural transformation, increasing credit access, promoting diversification of income sources, and increasing support from non-governmental organizations (Kumssa & Jones, 2010; Assan et al., 2018). Similarly, solutions to alleviate gender disparities within comparable communities include increased financial management, land and livestock rights, and exposure to grassroots organizations that focus on female empowerment for female household members (Khumalo & Freimund, 2014; Sharaunga et al., 2015; Meinzen-Dick et al., 2019; Goldman & Little, 2015). One solution implemented to concurrently address multiple issues facing rural communities is the institution of Community Based Conservation Areas (CBC's). CBC's are established with the purpose of economically supporting resident human populations while protecting natural areas and biodiversity (Berkes, 2007; Khumalo &

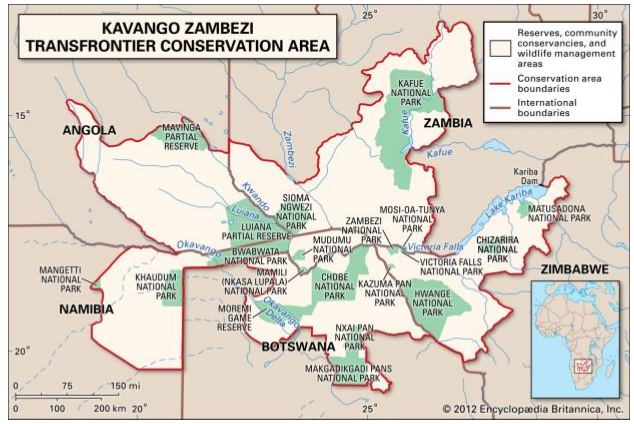
Freimund, 2014). Within this framework, local communities have agency over their natural resources to promote economic growth through sustainable activities, such as ecotourism (Khumalo & Freimund, 2014; Blaikie, 2006; Reibelt & Nowack, 2015). While CBC's are deemed highly effective in theory, it is difficult to quantify success because of the complexity of problems addressed and the prioritization of desired outcomes (Berkes, 2007; Reibelt & Nowack, 2015; Salerno et al., 2016). In addition, the outcomes of CBC's are often not monitored and evaluated for evidence of success (Blaikie, 2006; Lee & Bond, 2018; Hackel, 1999; Bajracharya et al., 2006; Lee & Bond, 2018; Mugisha & Jacobson, 2004). In particular, little empirical data is available on the impact of CBC's on climate vulnerability and gendered disparities.

My research compares gendered differences in indicators of climate vulnerability from CBC households of SSA to expected regional patterns. I start my study by reporting on expected changes and challenges of regional climate variation, regional gendered disparities, and a brief overview of CBC's. I then conduct a literature review of publications assessing gendered climate vulnerability in SSA. From this review, I compile a list of indicators of climate vulnerability to use comparatively in my data analysis. Next, I construct the framework of a Livelihood Vulnerability Index (LVI), comprised of eight major components that influence indicators found in my literature review and chosen to encapsulate the many dimensions of climate vulnerability. Following, I utilize data from the larger Kavango Zambezi Vulnerability and Adaption Project (KAZAVA) to assign each household an LVI score. The data provided through KAZAVA was collected through household livelihood surveys in Zambia, Botswana, and Namibia with the purpose of assessing the impact of CBC's on income, material wealth, agriculture, natural resource use, socio-demographic information, wildlife crop damage, and other variables. I then perform a series of inferential statistics to identify where gendered climate disparities exist within our focal communities. To support my study further, I next assess community employment by gender and the potential impacts of CBC participation on LVI and major component scores. I discuss the study's finding in comparison to results of regional studies from the literature review in order to identify potential implications of CBC's on gender disparities in the selected communities. I conclude my research with a discussion of study gaps and potential avenues for further research.

2. Study Background

2.A. Expected Changes and Challenges of Regional Climate

The study area, Kavango Zambezi Transfrontier Conservation Area (KAZA) is located in the countries of Namibia, Zambia, Botswana, Angola, and Zimbabwe (*Map 1*). These countries are part of a larger geographic area referred to as Southern Africa (*Map 2*). Southern Africa is in turn is part of a larger area of the African continent referred to as SSA. The human population of SSA is vastly composed of rural communities that account for 85% of the area's residents (Hoscilo et al., 2015). In addition, SSA also has reported to have the highest incidence of poverty in the world since 1990 (Milazzo & van de Walle, 2017).



Map 1: KAZA
Courtesy of Pallardy, 2012. Showing the borders of the Kavango Zambezi Transfrontier Area, national parks, and associated countries.

Map 2: Southern Africa
Courtesy of AFMI, 2020. Showing the countries in green that comprise the greater regional area of Southern Africa in green.

Southern Africa is classified as a semi-arid region with high interannual and seasonal rainfall variability that is characterized by pronounced droughts and floods (Mason & Jury, 1997). Natural variation in the climate of Southern Africa is influenced by interdecadal sea-surface temperature variability, the Agulhas

Current system, El Niño, and stratospheric oscillation of zonal winds. (Mason & Jury, 1997). However, there is a growing body of evidence indicating long-term climactic change including empirical data on decreases in mean average rainfall as well as increases in air temperature, interannual rainfall variability, severity of droughts, desertification, and frequency of extreme flood events (Mason & Jury, 1997; Connolly-Boutin & Smit, 2016). These assessments of extreme events are considered consistent with expected results from changes in the hydrological cycle due to climate change (Mason & Jury, 1997; Fauchereau et al., 2003; Hoscilo et al., 2015).

Extreme weather and climate events are expected to directly impact human populations throughout the Southern Africa region. Records of past extreme flooding events in February 2000 and September 1987 took the lives of 600 and 300 people respectively (Hoscilo et al., 2015). In addition, extreme heat events are expected to cause human deaths from heat stress, exhaustion, and stroke (Serdeczny et al., 2017). Finally, outbreaks of transmissible diseases, such as Rift Valley fever and malaria, often follow destruction caused by flooding and higher temperatures (Serdeczny et al., 2017).

Overall climactic shifts are predicted to gradually affect livelihoods in rural SSA. Higher temperatures are predicted to result in higher food prices and decreases in industrial growth, opportunities for job creation, and poverty reduction (Alagidede et al., 2016; Brooks, 2018). Notably, aggregated crop yields may decrease by up to 24% (Connolly-Boutin & Smit, 2016). It is important to note that there is also opportunity for job creation within SSA countries through investment in adaptive technologies, such as renewable energy (Brooks, 2018).

Climate related challenges are expected to manifest in the daily lives of communities and households in a multitude of ways. These include decreases in food security, ability to raise livestock, agricultural yields, water scarcity, and individual human health (Connolly-Boutin & Smit, 2016). Many households with land-dependent occupations, such as pastoralists, are pushed to adjust their livelihoods dramatically in order to adapt (Connolly-Boutin & Smit, 2016; Balehey et al., 2018; Serdeczny et al., 2017). Household migration is expected to be particularly high in SSA, posing uncertainty and risk to relocating climate refugees, as well as adding pressure to receiving areas (Serdeczny et al., 2017). Such transitions are experienced differently by each individual, and their full impacts on each human life cannot be fully or accurately encapsulated.

2.B. Regional Gender Disparities

Regional assessments have found SSA to have high gender inequality compared to other regions worldwide (OECD, 2014). Variables that lead to high gender-inequality include early marriage practices for underage girls, restricted autonomy of women over their physical bodies, skewed sex-ratios that indicate missing girls due to the bias of families towards sons, restricted access to own land and non-land assets, and restricted political voices (OECD, 2014). Challenges for women in SSA also occur in association with maternal and reproductive health, through low contraceptive use, low access to healthcare, and an unusually high maternal mortality ratio that accounted for two-thirds of all maternal deaths globally in 2015 (Pons-Duran et al., 2019).

Patterns of these regional gender-based issues are found to persist in countries encompassing the study area. For example, gender-based issues discussed in focus groups of the Caprivi Nation of Namibia include reports of higher incidence of HIV contraction among women, domestic violence, and lack of sexual freedom (Thomas, 2007). In Botswana, a Gender Affairs Department report published in 2014 found that women remain the most affected by poverty, unemployment, HIV and AIDS, experience high gender-based violence at a prevalence rate of 67% and comprise only 8.2% of national parliamentarians (GAD, 2014). Finally, in Zambia, the United Nation Development Programme (UNDP) reported that women earn 77 cents to a men's dollar for equal work, represent only 13% of agricultural landholders, comprise only 24% of national parliamentarians, and experience gender based violence at a prevalence of 35% (UNDP, 2020). These presented issues represent a small sample of topics of interest in regard to gender equality in these countries and are by no mean exhaustive of the disparities present.

It is also important to note that neither SSA or the separate countries of Namibia, Botswana, and Zambia are homogenous entities. A multitude of ethnicities and distinct cultural traditions are present within these areas. Ethnic groups present within our study group include Subiya, Lozi, Tawana, Thimbukushu, Fwe, San, and Yeyi. Currently, a gap in literature exists on specific cultural traditions and gendered practices of ethnic groups included in this study.

In addition to conventional gendered disparities, amassed attention has started to recognize the gendered dimensions of challenges associated with climate change. Women, who traditionally carry the burden of water duties, must travel longer distances to retrieve water during times of drought and scarcity (Balehey et al., 2018; Assan et al., 2018). In addition, food scarcity is often amplified in female household members. Societal standards dictate that woman and girls must often give up food for the benefit of male members of the family

(Balehey et al., 2018). Mother's may also feel compelled to go hungry in order to feed their children (Balehey et al., 2018).

Although my research focuses more on indirect impacts of climate change on households, it is also important to note that further gendered vulnerabilities occur during acute shocks of natural disasters directly caused by climate change. Woman experience increased risk of death, violence, and sexual assault during and after climate change-induced disasters (Nelleman & Hislop, 2011; Balehey et al., 2018).

Regional gender inequalities and vulnerabilities manifest in household headship through a multitude of pathways. FHH have been found to hold less wealth then MHH and are attributed with the challenges of a "triple burden" of female labor market disadvantages, time constraints of balancing household tasks with employment, and the lack of a partner who is also earning wages (Rogan, 2013; Nwosu & Ndinda, 2018). Additionally, as stated before, literature supports that rural FHH are exceptionally susceptible to the shocks of climate change when compared to MHH households due to decreased diversification of livelihood strategies, limited access to social safety nets such as credit, limited ownership of wealth and assets, among other factors (Connolly-Boutin & Smit, 2016; Alhassan et al., 2019; Balehey et al., 2018; Assan et al., 2018; Ibrahim et al., 2019; Tibesigwa & Visser, 2016). However, it is important to note that some ideas associated with lower wealth and resilience are starting to be challenged by recent studies that find no significant differences in FHH wealth when compared to MHH (Fuller & Lain, 2019; Milazzo & van de Walle, 2017; Baffoe & Matsuda, 2018; Balikoowa et al., 2018). My research will continue to explore the connection between female headship and climate vulnerabilities, as well as the role of CBC's in providing solutions to persisting gender inequalities through increased livelihood diversification, income for women, and other opportunities for empowerment. In particular, the relationship of gender to climate change will be covered extensively in the literature review.

2.C. Overview of Community Based Conservation Areas

CBC areas are defined as "wildlife conservation efforts that involve rural people as an integral part of a wildlife conservation policy", featuring key elements such as local community participation in the management and planning of resource use (Hackel, 1999). CBCs are often referred to as a supposed "panacea" in literature (Berkes, 2007; Reibelt & Nowack, 2015), a word synonymous with a supposed "cure-all". CBC's were popularized in the 1980's in order to balance human resource needs with the preservation of biodiversity (Reibelt & Nowack, 2015; Hackel, 1999). This solution was created in order to address the increasing exclusion of rural people from conservation efforts (Hackel, 1999). This exclusion is deemed problematic because of historical colonialist roots, negative impacts on local economies, and restriction of natural resources integral to community survival (Hackel, 1999).

Although CBC's are a highly touted solution in theory, literature often notes the lack of studies and empirical data to back such claims up (Reibelt & Nowack, 2015; Hackel, 1999; Bajracharya et al., 2006; Lee & Bond, 2018; Mugisha & Jacobson, 2004). Examples of variables not always considered when discussing conceptual benefits of CBC's include restricted revenue generation, limitations of land use, and population growth (Hackel, 1999). Many completed studies conclude that CBC's must be instituted concurrently with other approaches in order to achieve desired results (Reibelt & Nowack, 2015; Bajracharya et al., 2006; Lee & Bond, 2018; Mugisha & Jacobson, 2004; Salerno et al., 2016).

Completed studies suggest that CBC's contribute to a mixture of positive and negative results. Negative results include heightened food insecurity due to increased human-wildlife conflict (Salerno et al., 2016), poor attitudes towards conservation areas that may lead to higher potential of conflict (Larson et al., 2016), and high expectations of results by locals that may lead to retracted support (Nanang & Nunifu, 2010). Positive results include increased household well-being (Salerno et al., 2016), ecological success characterized by significantly higher densities of resident wildlife (Lee & Bond, 2018), use of sensitive natural resources that are both sustainable and profitable (Aheto et al., 2016), improvements in access to forest resources, basic infrastructure development, health, sanitation, and social services (Bajracharya et al., 2006), and mitigation of specific threats to conservation efforts (Mugisha & Jacobson, 2004).

Other particularly relevant studies note CBC and ecotourism employment opportunities show positive impacts on indicators of gender equality. Conservancy related jobs help to close income disparities between male and female households, which contributes to heightened self-esteem and status in women, decreased domestic work burden, and inadvertently contributes to higher awareness of health education (Khatiwada & Silva, 2015). A study performed in the Uibasen and Mayuni Conservancies in Namibia found that FHHs are statistically economically better off with tourism employment than those without it, while MHHs demonstrated no differences (Khatiwada & Silva, 2015). A separate study conducted in the Kwandu Conservancy of Namibia found that the conservancy has created new employment opportunities for women, contributed to more diversified livelihoods for women, expanded female social networks and subsequently the informal exchange of human rights information, and challenged the patriarchal order by putting women in management positions (Khumalo & Freimund, 2014).

While the previously listed studies have begun to explore the impacts of CBC's on gender in regard to food insecurity, human-wildlife conflict, local perception, household well-being, and wealth, no study has yet assessed the interaction of gender and multi-dimensional climate vulnerability with considerations for the unique CBC setting. My research strives to serve as a preliminary assessment for potential impacts of CBCs on gendered climate vulnerability and provide a framework of methodology for comparative assessment in future studies.

3. Literature Review

I conducted a literature review by searching for studies in SSA that had a strong focus on the relationship between gender and climate. Here, gender is defined in three ways. First, gender is considered as the sex of an individual participant. Second, gender is representative of greater household measurements through the gender of the household head, testing FHH against MHH. Finally, gender of the household head is considered along with closer scrutiny of partnership status. These household groupings are defined as female de jure (FDJ), female de facto (FDF), male with a spouse (MSP) and male without a spouse (MNS). FDJ households are defined as those in which the female head of house is not attached to a partner, as a result of being widowed, divorced, or never married (Rogan, 2013). FDF household heads are married, but not living with a husband or partner, so may be standing in as the household head with additional support. FDJ households illicit additional attention over FDF households, as they are often found as the most vulnerable group in reviewed literature (Tibesigwa & Visser, 2016; Rogan, 2013; Flato et al., 2017).

The literature review was conducted informally utilizing the database Web of Science. The search produced 22 studies related to gender and climate in countries of SSA. While I was thorough in searching for relevant literature, the search was not exhaustive and other literature may persist outside of what was encapsulated in my study. The studies evaluated are listed in Table 1 (*pg. 16)* along with details regarding study location, variables assessed, type of research conducted, and concluded vulnerable groups. The majority of studies (n = 20) state that FHH demonstrate more climate related vulnerability. These increased vulnerabilities are attributed to food insecurity, productivity gaps, lifestyle diversification, adaption measures, social treatment, wealth, awareness of hazards, along with a number of other indicators. However, it is important to note that two contradictory studies reviewed found higher climate vulnerability in MHH. These studies, located in Ghana and Uganda, attribute this trend to higher reliance of MHH on agriculture and overlooked vulnerability of MNS households when compared to other household types, respectively (Baffoe & Matsuda, 2018; Balikoowa et al., 2018).

This literature review informed my decision of how to characterize climate vulnerability within the studied KAZA community. Utilizing framework developed by Hahn et al. (2009), I compiled a LVI comprised of eight major components that encapsulate the many dimensions of climate change. The LVI and associated major components are described extensively in the Methods section (*pg. 21*) and illustrated in Figure 1 (*pg. 23*). Table 2 (*pg. 19)* describes empirical data found relating to each of the chosen major components to provide a basis for expected results, and a comparison for tests performed later between different groupings of MHH and FHH. Much of the data listed in Table 2 is rearranged from the studies listed in Table 1. However, a few studies that did not explicitly state a connection to climate were added to the assessment because of their role in evaluating the relationship of gender in SSA and variables relevant to the LVI (Rogan, 2013; Dungumaro, 2008).

Table 1: Studies on Gendered Climate Vulnerability in SSA						
#	Title	Citation	Location	Variables of Interest	Research	Results
1.	Effects of climate change and livelihood diversification on the gendered productivity gap in Northern Ghana	(Adzawla et al., 2019)	Ghana	Productivity gap and livelihood diversification		
2.	Gender dimension of vulnerability to climate change and variability: Empirical evidence of smallholder farming households in Ghana	(Alhassan et al., 2019)	Ghana	LVI compiled of socio-demographic profile, social networks, water, food, health, livelihood strategies, climate variability		
3.	Coping with and Adapting to climate Change: A Gender Perspective from Smallholder Farming in Ghana	(Assan et al., 2018)	Ghana	Coping and adaption measures adopted by household heads		
4.	An Empirical Assessment of Household Livelihood Vulnerability: The Case of Rural Ghana	(Baffoe & Matsuda, 2018)	Ghana	LVI compiled of socio-demographic profile, livelihood strategies, health, food, water, institutional influence, and climate variability		
5.	Traditional gender inequalities limit pastoral women's opportunities for adaptation to climate change: Evidence from the Afar pastoralists of Ethiopia	(Balehey et al., 2018)	Ethiopia	Social treatment, wealth ownership, livestock inheritance, income and expenditure, household decision making, community participation, food security, adaption strategies		
6.	Gender stereotyping: Evidence from gender differentiated household vulnerability to climate change in Eastern Uganda	(Balikoowa et al., 2018)	Uganda	Vulnerability index composed of livelihood strategies, social networks, income, food security, land characteristics, and natural disaster exposure		

7.	Integrating gender into index-based agricultural insurance: a focus on South Africa	(Born et al., 2019)	South Africa	Benefit of index insurance for climate risk management		
9.	Determinants of farmers' choice of adaptation methods to climate change in the Nile Basin of Ethiopia	(Deressa et al., 2009)	Ethiopia	Climate change adaption practices		
10.	Women, Weather, and Woes: The Triangular Dynamics of Female-Headed Households, Economic Vulnerability, and Climate Variability in South Africa	(Flato et al., 2017)	South Africa	Fluctuations of household income related to rainfall		
11.	Are female-headed households less resilient? Evidence from Oxfam's impact evaluations	(Fuller & Lain, 2019)	Mali, Niger, Ethiopia, Kenya, Zambia, Chad	Resilience, measured through indicators similar to LVI approach		
13.	Vulnerability to recurrent shocks and disparities in gendered livelihood diversification in remote areas of Nigeria	(Ibrahim et al., 2019)	Nigeria	Vulnerability to natural hazard-induced and cattle rustling-driven shocks		
14.	Gender vulnerability to climate variability and household food insecurity	(Kakota et al., 2011)	Malawi	Food insecurity, coping strategies, and livelihood diversification		
15.	Access to energy sources in the face of climate change: Challenges faced by women in rural communities	(Ketlhoilwe & Kanene, 2018)	Botswana	Gendered challenges of energy access and sources		
16.	Diagnosing Climate Adaptation Constraints in Rural Subsistence Farming Systems in Cameroon: Gender and Institutional Perspectives	(Nchu et al., 2019)	Cameroon	Land tenure, gender inequality and resilience of farming systems		

17.	Climate change, household vulnerability and smart agriculture: The case of two South African provinces	(Ncube et al., 2016)	South Africa	Impact of climate change on agricultural productivity and food insecurity		
18.	Gender, social capital and adaptive capacity to climate variability: A case of pastoralists in arid and semi-arid regions in Kenya	(Omolo & Mafongoya, 2019)	Kenya	Perceptions of vulnerability within the community		
19.	Gendered perception and vulnerability to climate change in urban slum communities in Accra, Ghana	(Owusu et al., 2019)	Ghana	Awareness of environmental hazards posed by climate change		
20.	Managing risk, changing aspirations and household dynamics: Implications for wellbeing and adaptation in semi-arid Africa and India	(Rao et al., 2020)	Ethiopia, Kenya, Mali, Ghana	Accessibility of risk-management strategies and subjective wellbeing		
21.	Assessing Gender Inequality in Food Security among Small-holder Farm Households in urban and rural South Africa	(Tibesigwa & Visser, 2016)	South Africa	Food security		
22.	Gender Differences in Climate Change Risk, Food Security, and Adaptation: A Study of Rural Households' Reliance on Agriculture and Natural Resources to Sustain Livelihoods	(Tibesigwa et al., 2015)	SSA regional study	Food security and natural-resource and agricultural reliance		

Table 1 Key:

Research:

Primarily Literature Review

Data Collected

Used Existing Data

Results:

More vulnerability observed in female groups

More vulnerability observed in male groups

Table 2: Empirical Support for Increased Female Vulnerability in each Major Component

Major Component	Study Metric	Country	Assessed Groups	Reference
Overall Climate Vulnerability	Compiled LVI Score	Ghana	FHH > MHH	(Alhassan et al., 2019)
		South Africa	FHH > MHH	(Ncube et al., 2016)
	Community perception of climate vulnerability	Kenya	FHH > MHH	(Omolo & Mafongoya, 2019)
Food	Food Major Component	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Food Insecurity	South Africa	FHH > MHH	(Dungumaro, 2008)
		South Africa	FHH > MHH	(Tibesigwa & Visser, 2016)
		Ghana	FHH > MHH	(Alhassan et al., 2019)
		Namibia	FHH > MHH	(Rao et al., 2020)
		SSA	FHH > MHH	(Tibesigwa & Visser, 2016)
	Dependence on Family Farm for Food	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Crop Diversity Index	Ghana	FHH > MHH	(Alhassan et al., 2019)
Wealth	Amount of wealth (personal)	Ethiopia	FP > MP	(Balehey et al., 2018)
	Amount of wealth (household)	South Africa	FHH > MHH	(Dungumaro, 2008)
		South Africa	FHH > MHH	(Tibesigwa & Visser, 2016)
	Household expenditure devoted to individuals	Ethiopia	FP > MP	(Balehey et al., 2018)
	Housing Conditions	South Africa	FHH > MHH	(Dungumaro, 2008)
	Household Income	SSA	FHH > MHH	(Rogan, 2013)
	Household Consumption	South Africa	FHH > MHH	(Tibesigwa & Visser, 2016)
Livelihood Diversification	Livelihood Strategies Major Component	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Agricultural Reliance	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Agricultural and Natural resource reliance	SSA	FDJ> FDF > MHH	(Tibesigwa et al., 2015)
	Farmland Ownership	Ghana	FHH > MHH	(Alhassan et al., 2019)
		Cameroon	FHH > MHH	(Nchu et al., 2019)
	Agricultural Livelihood Diversification Index	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Members working outside of the community	Ghana	FHH > MHH	(Alhassan et al., 2019)

	Productivity of Farmland	Ghana	FHH > MHH	(Adzawla et al., 2019)
	Livelihood Diversification	Ghana	FHH > MHH	(Adzawla et al., 2019)
	Natural Resource Access	Cameroon	FHH > MHH	(Nchu et al., 2019)
Socio-Demographic	Household Size	South Africa	FHH > MHH	(Dungumaro, 2008)
		SSA	FHH > MHH	(Rogan, 2013)
	Proportion of Dependents	South Africa	FHH > MHH	(Dungumaro, 2008)
		SSA	FHH > MHH	(Rogan, 2013)
	Number of Orphans	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Education of Household Head	South Africa	FHH > MHH	(Dungumaro, 2008)
		South Africa	FHH > MHH	(Tibesigwa & Visser, 2016)
Water	Water Major Component	Ghana	MHH, FHH	(Alhassan et al., 2019)
	Water Source	South Africa	FHH > MHH	(Dungumaro, 2008)
		Ghana	FHH > MHH	(Alhassan et al., 2019)
	Reported water conflicts	Ghana	FHH > MHH	(Balehey et al., 2018)
Health	Chronic illness in household	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Body Mass	Ethiopia	FP > MP	Balehey et al., 2018)
	Social Networks Major Component	Ghana	FHH > MHH	(Alhassan et al., 2019)
	Life Satisfaction	Namibia	FHH > MHH	(Rao et al., 2020)
Social Networks	Household level decision power	Ethiopia	FP > MP	(Balehey et al., 2018)
	Community level decision power	Ethiopia	FP > MP	(Balehey et al., 2018)
	Reliance on Social Grant Income	SSA	FHH > MHH	(Rogan, 2013)
	Access to Local Government Assistance	Ghana	FHH > MHH	(Alhassan et al., 2019)
Environmental Shock	Vulnerability to cattle rustling	Nigeria	FP > MP	(Ibrahim et al., 2019)
	Vulnerability to climate-change induced challenges	Ethiopia	FP > MP	(Balehey et al., 2018)
		Ghana	FP > MP	(Owusu et al., 2019)
	Climate Shocks	South Africa,	FHH > MHH	(Flato et al., 2017)
	Response to climate change through adaptive activities	Ethiopia	FHH > MHH	(Deressa et al., 2009)
		Ghana	FHH > MHH	(Assan et al., 2018)
	Education and awareness of Climate Change	Ghana	FP > MP	(Owusu et al., 2019)

4. Methodology

4.A. Description of study area

This study took place within the KAZA (*MAP 1*). Member nations of KAZA include Angola, Botswana, Namibia, Zambia and Zimbabwe. Formally established in 2011, with the purpose of sustainably managing the ecosystem and cultural resources using the best conservation models possible, KAZA is an important tourist destination and Africa's largest remaining wilderness (KAZA TFCA, 2019). The KAZA landscape is described as largely semi-arid dry forest and savanna with extensive wetlands and shrublands. This area is prone to drought and flooding due to its highly variable wet and dry seasons. As discussed previously, shifts in regional climate and rainfall patterns are likely to occur as a result of climate change.

Human-nature dynamics within KAZA range from national parks to urban areas. Thirty-six protected areas are established with varying levels of regulation. National parks have strict regulations on human activities including tourism and natural resource collection. Other areas harbor resource extraction and farming, which lends to higher ecosystem impacts.

Site selection for data collection within the greater KAZA area was conducted purposefully in consultation with the KAZA secretariat, traditional authorities, and collaborating institutions. The chosen CBC's are the Chobe Enclave Conservation Trust (CECT) in Botswana; Mashi Conservancy in Namibia; and Lower West Zambezi Game Management Area (LWZ GMA), in Zambia.

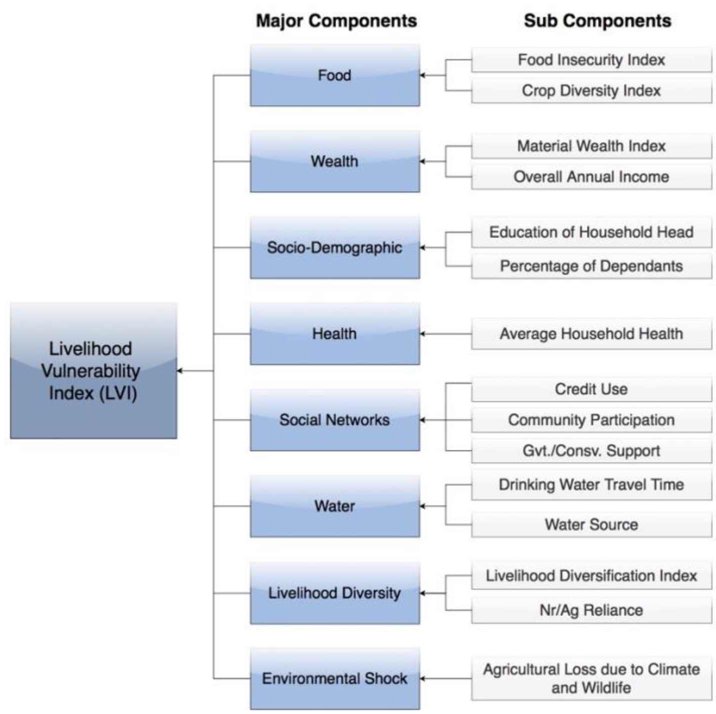
4.B. Data collection process

This study is integrated into the larger KAZAVA supported by the National Science Foundation (NSF). Data for this larger project was collected through surveys conducted with in two languages, Setswana in Botswana and Lozi in Namibia and Zambia. Forty-eight households were selected in each of five villages chosen for a total of 720 households. The first household was chosen at random. The following households were chosen by dividing the number of households in the village by 48 to result in the number *n.* Afterwards, the *nth* household counted after a previously surveyed household was surveyed afterwards. Surveys were administered by trained enumerators, who refrained from conducting surveys within their home villages. Surveys were conducted orally and proceeded by verbal consent with the stated purpose of obtaining information about the livelihood of households, and what resources they use in the landscape. Contents of the survey were concentrated into 5 sections; Human roster/human capital, health and nutrition, financial capital, social capital, and natural capital. This study utilizes data from the entire study set of households (*n = 720*) and a subset of questions from all the 5 sections. Field work was conducted between May and July, in the dry seasons of 2017 and 2018.

4.C. Overview of Livelihood Vulnerability Index

Vulnerability can be assessed through a variety of methods, as discussed previously in the literature review and listed in Table 1 (*pg. 16*). For this study, I chose to characterize vulnerability using the LVI developed by Hahn (2009). I chose this method because of its ability to incorporate many dimensions of vulnerability and for its optimal utilization of information from the provided dataset. The LVI method is an indicator approach to vulnerability, which is favorable over econometric approaches due to its computational accessibility and appeal to readers (Hahn et al., 2009). I have adapted Hahn's methods to apply an LVI score to each household. This approach was chosen in order to group and assess differences by household head as well as by country. All calculations were performed using R-Studio. Appendix C (pg. 75) lists all code used for this project.

Figure 1: Components of the Livelihood Vulnerability Index



The LVI is made up of several major components, which are combined (*Eq. 5, pg. 25*) to yield an overall LVI score assigned to each household. The 8 major components chosen for this study are food, household wealth, livelihood diversity, socio-demographics, water availability or access, health, social networks, and environmental shock. Each major component is made up of 1-3 sub-components calculated from the survey data and chosen to characterize the major components (*Fig. 1*).

The subcomponents were first calculated in different scales and then standardized utilizing *Equation (1)*:

$$index_{S_H} = \frac{S_H - S_{min}}{S_{max} - S_{min}} \tag{1}$$

Where S_H is the original sub-component for each respective household and S_{min} and S_{max} are the minimum and maximum values, respectively, for each sub-component using data from the entire dataset. Before standardization, the distribution of each subcomponent was assessed, and natural log transformed as needed using *Eq. (2)*:

$$l_{S_H} = ln\,(S_H + 1) \tag{2}$$

Where S_H is the original sub-component for each respective household and is l the log transformed value.

Additionally, it was necessary to maintain that each sub-component indicated higher vulnerability with a higher value, and lower vulnerability for a lower value. Many variables exhibited this pattern and did not need to be transformed further, such as percentage of dependents, drinking water travel time, and calculated indices such as the Crop Diversity Index and the Livelihood Diversification Index. Other sub-components did not demonstrate this relationship, such as overall annual income, community participation, and level of education of the household head. The inverse of these sub-components was taken using *Eq. (3)* and utilized in place of the original calculated values:

$$Inv_{S_H} = \frac{1}{S_H + 1} \tag{3}$$

Where S_H is the original sub-component for each respective household and is Inv the inverse of this value.

The standardized and sometimes log-transformed and inverted sub-components were used to find the value of each major component using *Eq. (4)*:

$$M_H = \frac{\sum_{i=1}^{n} index_{S_H i}}{n} \tag{4}$$

Where M_H represents one of the eight major component scores for each respective household [Food Component (FC), Wealth Component (HWC), Livelihood Diversification Component (LDC), Socio-Demographic Component (SDC), Water Component (WC), Health Component (HC) Social Networks Component (SNC), or Environmental Shock Component [ESC], $index_{S_H i}$ represents the sub-components, indexed by i, that make up each major component, and n is the total number of sub-components in each major component.

Finally, the LVI was calculated using *Eq. (5), also expressed as* also expressed as *Eq. (6)*:

$$LVI_H = \frac{\sum_{i=1}^{8} W_{M_i} M_{Hi}}{\sum_{i=1}^{8} W_{M_i}} \tag{5}$$

$$LVI_H = \frac{w_{FC}FC_H + w_{HWC}HWC_H + w_{LDC}LDC_H + w_{SDC}SDC_H + w_{WC}WC_H + w_{HC}HC_H + w_{SNC}SNC_H + w_{ESC}ESC_H}{w_{FC} + w_{HWC} + w_{LDC} + w_{SDC} + w_{WC} + w_{HC} + w_{SNC} + w_{SNC}} \tag{6}$$

Where LVI_H, the LVI score for household H, equals the weighted average of the eight major components. Weights, w_{Mi}, are equal to the number of subcomponents in each major component, which ensures equal contribution of each sub-component to the overall index. This equal weighting approach, developed by Sullivan et al., 2002, and utilized by Hahn et al. 2009, is preferred to eliminate bias in weighting on part of the researcher. A step by step walk through on how to calculate a LVI, as used in my research, is detailed clearly in Appendix A of Hahn et al., (2009). In the next subsection, I walk through each of major components and the processes used to calculate the respective sub-components.

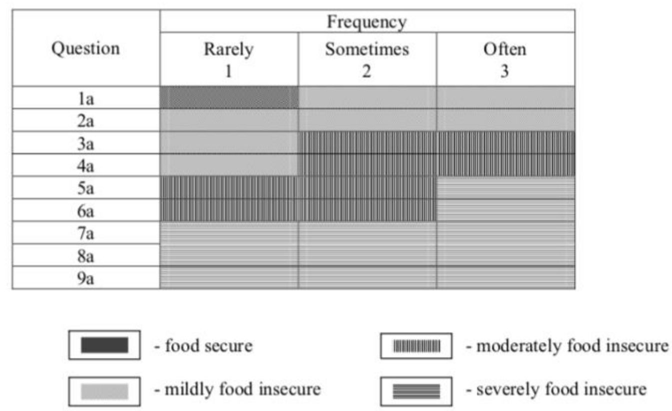
4.D. Calculation of Sub-Components

4.F.i. Food

The Food major component (FC) is composed of two sub-components; food insecurity and crop diversity index.

Food insecurity was calculated using Household Food Insecurity Access Scale (HFIAS) protocol developed by the United States Agency for International Development (USAID) (Coates et al., 2007). Survey questions used to determine food insecurity are listed in *Appendix A* (*pg. 71)*.

Table 4: HFIAS Categories of Food Insecurity



Courtesy of (SOURCE). This classification table was used in conjuncture with survey questions listed in Appendix A (pg. 71) to determine the household level of food insecurity.

The questions were designed with the intention to evaluate the occurrence and frequency-of-occurrence of various physical and psychological factors related to household food supply. Based on a set of rules applied to ordinal question answers, each household was assigned a category of food insecurity (*Table 4*). The highest categorical rating of insecurity a household's answers qualified for was applied as a score for the food insecurity subcomponent.

Crop diversity index was calculated using methods adapted from the World Bank, 1997. Survey questions asked households if crops from seven categories were grown in an average year. The seven categories were: maize, sorghum, millet, beans and peanuts, pumpkins and melons, sweet reed, and other. The index was calculated as the inverse of the number crop categories grown by a house plus one, demonstrated in *Eq. (7)*:

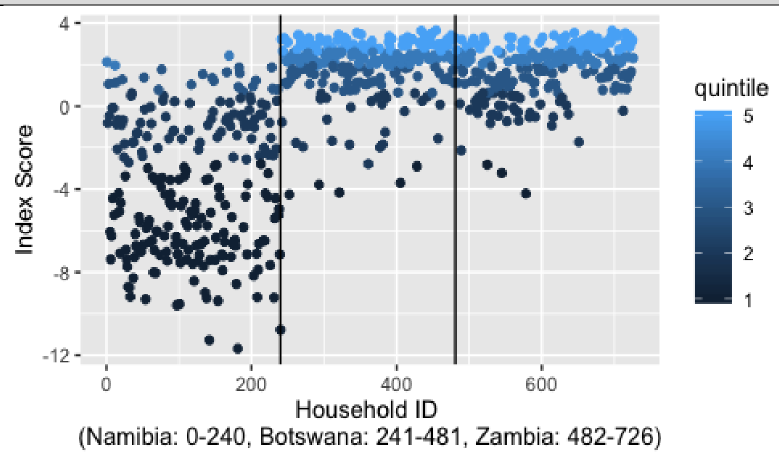
$$CDI_H = \frac{1}{n_{C_H} + 1} \tag{7}$$

Where n_{C_H} is the number of crop categories grown by the respective household and CDI_H is the crop diversity index score for that household. Each household was assigned a score from 0-1 using this formula. Households who grew no crops were assigned the highest score of 1, while households who had very high crop diversity were assigned a score closer to 0. This index score was used directly as the crop diversity index subcomponent.

4.F.ii. Wealth

The Wealth Major component (HWC) is composed of two sub-components; material wealth index and overall annual income.

Figure 2: Material Wealth Index



The material wealth index ranks each household by relative wealth, with 1 indicating high wealth and 5 indicating low wealth.

The material wealth index (*Figure 2*) was calculated using Demographic and Health Survey program (DHS) protocol developed by USAID (DHS Program, 2020). This method utilizes principle component analysis (PCA) to determine which combination of household assets most accurately characterize household wealth within the studied community. This process is important to prevent misconstruing the ownership of an item (such as a hoe) as an indicator of wealth, when it may actually signify the absence of an item (such as a plow) that more accurately depicts household wealth (DHS Program, 2020). This method is preferred over an additive method, in which each item is attributed a score depicting its worth, because of the difficulties associated with unbiased assignments of worth. The DHS protocol used outlines steps necessary to transform survey data into the correct format before running the PCA.

The scores of the first principle component are assigned to each household and classified into quintiles. Households with a score of 1 demonstrate the greatest relative household material wealth, while those with a score of 5 demonstrate the least relative household wealth. This index score was used directly as the material wealth index subcomponent. Overall annual income was calculated as a sum of all cash income, agricultural cash value, and natural resource cash value using *Eq. (8)***:**

$$OAI_H = CI_H + AGV_H + NRV_H \tag{8}$$

Where OAI_H represents the overall annual income for the household, CI is cash income, AGV is agricultural cash value, and NRV natural resource cash value.

This method was chosen to prevent the exclusion of households predominately reliant on agriculture and natural resources. Vulnerabilities associated with heavy agricultural and natural resource reliance are later accounted for under the Livelihood Strategies major component. All variables used to calculate overall income are listed in *Table 4*. All income values were converted into US dollars using an exchange rate of 10 for the Botswana pula and 13 for the Namibia dollar (used for Namibia and Zambia).

Table 4: Overall Annual Income ***Calculated per Household***		
Cash Income	*Agriculture*	*Natural Resources*
Tourism	Maize*	Firewood*
Shop/office work	Sorghum*	Thatching grass*
Non-tourism	Beans and Peanuts*	Fish*
Small-business or crafts	Pumpkins and Melons*	Papyrus*
Other sources	Sweet Reed*	Grapple*
Remittances	Other crops*	Building poles*
Labor program	Cattle Sales	Mud*
Other cash payments	Goat Sales	Palm leaves*
	Sheep Sales	Water lily*
	Donkey Sales	Medicinal Plants*
	Horse Sales	Edible plants*
	Poultry Sales	Birds and animals*
	Pig Sales	Other natural resources*

**Indicates the cash value of the total amount of product harvested*

Cash income was defined using protocol from the United States International Revenue Service (IRS) and therefore includes wages, remittances, and other sources, but does not include credit (IRS, 2020). The cash income total did not include government and conservancy aid, as this total was used as a separate sub-component under the Social Networks major component. Agricultural and natural resource cash values were calculated by multiplying all agricultural and resource products harvested by the household by country-specific product prices, demonstrated in *Eq. (9):*

$$AGV_H = (amount_{P1_H} \times price_{P1_C}) + (amount_{P2_H} \times price_{P2_C}) \ldots n_P \tag{9}$$

Where AGV_H represents the total agricultural cash value per household, $amount_{P1_H}$ Is the amount of product 1 ($P1$) harvested by the household (H) and $price_{P1_C}$ is the country specific product price for product 1. This repeats for product 2, and so on. This process is replicated todetermine AGV_H.

Table 5: Country Specific Product Prices *Listed as price per unit (differs by crop)*				
Resource	***Product***	***Botswana***	***Namibia***	***Zambia***
Agriculture	Maize	1.17	1.00	0.83
	Sorghum	1.350	0.985	0.620
	Millet	0.74	0.74	0.74
	Beans/Peanuts	5.05	8.08	2.09
	Pumpkins/Melons	7.24	7.24	7.24
	Sweet Reed	0.86	0.86	0.86
Natural Resource	Firewood	1.14	0.73	1.55
	Grass	1.17	0.68	0.31
	Fish	0.78	0.77	0.21
	Papyrus	3.08	3.08	3.08
	Grapple	1.175	1.710	0.640
	Reeds	5.06	1.68	0.69
	Poles	1.34	1.21	1.41
	Mud	0.5	0.5	0.5
	Palm	2.32	7.69	0.20
	Lily	1.04	1.04	1.04
	Medicinal Plants	0.50	0.35	0.20
	Edible Plants	0.55	0.22	0.05
	Other (NR)	0.2	0.2	0.2

The values for all agricultural and natural resource products were summed to yield the total cash value, which was then combined with cash income for each household [*Eq. (8), pg. 28]*. Each country-specific product price was found by dividing the country average total earned from product sold by the country average amount of product sold, demonstrated in *Eq. (10):*

$$Price_{X_C} = \frac{Average\ of\ x\ sold\ _c}{Average\ money\ made\ from\ x\ _c} \tag{10}$$

Where x represents the product and c represents the country. If a product price wasn't available for a country, the average of the other countries, or one country, was taken as the price. Product prices were assessed and adjusted by a researcher present during data collection to eliminate badly skewed prices from insufficient data. The final prices per product are listed in Table 5**.** This method was chosen in lieu of reported amount earned from each product in order to accurately credit amounts of each product used by the household or traded in addition to amount sold.

Calculation of agricultural values did not include the cash value of each livestock animal. This is in part to avoid redundancy, as livestock counts were incorporated into the material wealth index. In addition, it was difficult to quantify the monetary inputs and outputs of each animal, and to encompass the uncertainty of eventual payout. Instead, money earned from livestock sales was directly added with agricultural cash values before being incorporated into overall annual income. The sum of cash income, agricultural cash value, and natural resource cash value were log transformed using *Eq. (2) (pg. 24)* and inverted using *Eq. (3) (page 24)* before being applied as the overall annual income sub-component.

Table 6: Education of Household Head

Value	*Description*
1	No education
2	Pre-school
3	Primary school
4	Junior secondary school
5	Senior secondary school
6	Technical college
7	University

4.F.iii. Socio-Demographic

The Socio-Demographic major component (SDC) is composed of two sub-components - education of household head and percentage of dependents.

Education of household head was reported directly by survey responders. Possible ordinal responses (*Table 6*) ranged from 1-7, with 1 indicating no education and 7 indicating the highest level of education. These responses were inverted using *Eq. (3)* (*pg. 24*) and then applied as the education of household head sub-component.

Percentage of dependents represents the proportion of the household reliant on support by other members of the household. This variable was calculated by dividing the total number of reported residents by a household sum of residents aged 17 and under and residents aged 65 and older. It is important to note that up to 12 residents could be reported in the household survey, and any people above this number were unaccounted for. The household proportion of dependents generated by this calculation (in decimal form) was directly applied as the percentage of dependents sub-component.

4.F.iv. Health

The Health major component (HC) is composed of the variable, average household health. Average household health was calculated by dividing the sum of all resident health status scores by the number of residents. Resident health status scores ranged from 1-3. These scores were defined to survey participants as 1: "Always able to work, or attend school", 2: "Usually able to work, or attend school", and 3: "Never able to work, or attend school". These scores were either self-reported or decided by the representative household member filling out the survey. This calculation was directly applied as the average household sub-component.

4.F.v. Social Networks

The Social Networks major component (SNC) was composed of three sub-components - credit use, government and conservancy support, and community participation.

Credit use was calculated by totaling household responses in US dollars of money received as informal loans, formal loans, and loans received from friends or family. The final total was log transformed using *Eq. (2)* (*pg. 24*) and inverted using *Eq. (3)* (*pg. 24*) before being applied as the overall credit use subcomponent.

Government and conservancy support were calculated by totaling household responses in US dollars of money received from government pension, government orphans and vulnerable children (OVC) support, government relief programs, and conservancy trust payments. The final total was log transformed using *Eq. (2)* (*pg. 24*) and inverted using *Eq. (3)* (*pg. 24*) before being applied as the overall government and conservancy support subcomponent.

Participation in community institutions was calculated by estimating how many community institutions the household is involved in. The eight categories of optional involvement were: kgotla (traditional or customary court), church, village development committee (VDC), parent teacher association (PTA), farmer's committee, crime prevention committee, co-op of non-governmental organization (NGO), and other. Participation in the conservancy was omitted from this category so it could be used later in correlation tests to each of the major components and overall LVI. The final sum of institutions for each household was inverted using *Eq. (3)* (*pg. 24*) before being applied as the overall participation in community institution subcomponent.

4.F.vi. Water

The Water major component (WC) is composed of two sub-components -drinking water travel time and water source.

Drinking water travel time was calculated as an average of reported household wet and dry season travel times. This average was log-transformed using *Eq. (2)* (*pg. 24*) and then applied as the drinking water travel time sub-component.

Water source was calculated by ordinally categorizing answers for reported household water sources. Categories ranged from 1 to 3, 1 being a private source, 2 being a communal closed source (including a community standpipe, well point, borehole, and neighbors' source), and 3 being an open river. This ordinal score was applied directly as the water source sub-component.

4.F.vii. Livelihood Diversity

The Livelihood Diversity major component (LDC) is composed of two variables - livelihood diversity index and natural resource and agricultural reliance.

Livelihood diversity index was calculated using the number of livelihood-sustaining activities each household participated in. Ten livelihood-supporting activities were included in this assessment: tourism work, shop/office work, casual wage work, small business or craft work, other income-generating work, participation in labor programs, receipt of remittances, participation in crop growing, participation in keeping livestock, and participation in gathering natural resources. The index was calculated as the inverse of the number of livelihood categories participated in by a house plus one, demonstrated in *Eq. (11):*

$$LDI_H = \frac{1}{n_{L_H} + 1} \tag{11}$$

Where n_{L_H} is the number of livelihood categories participated in by the respective household and LDI_H is the crop diversity index score for that household. Each household was assigned a score from 0-1 using this formula. Households who demonstrated high livelihood diversity were assigned scores close to 0, while households who had very low livelihood diversity were assigned a score closer to 1. This index score was used directly as the crop diversity index subcomponent.

Natural resource and agricultural reliance were calculated as the percentage (in decimal form) of household income derived from natural resource and agricultural products (*Table 4, pg. 29*). In this instance, household income was modified from overall annual income to include government and conservancy support. This value was used directly as the natural resource and agricultural reliance subcomponent.

4.F.viii. Environmental Shock

The Environmental Shock major component (ESC) is composed of one variable - agriculture lost to wildlife and climate.

Agriculture lost to wildlife and climate was calculated as a sum of reported products lost multiplied by respective country-specific product prices. Calculations for country-specific product prices, listed in US dollars (*Table 5, pg. 30*), are detailed in calculation of the overall annual income sub-component. This total was log transformed using *Eq. (2)* (*pg. 24*) before being applied as the agriculture lost sub-component.

4.F.ix. Discarded Sub-Components

Other variables were considered for sub-components based on the literature review and data availability. These included number of orphans (SDC) and livestock water distance (WC). A sub-component describing the number of orphans per household was discarded because it had too little variation to be deemed useful. A sub-component listing livestock water distance was considered too reliant on other factors to be used as a stand-alone variable. For instance, some households may have a very long travel time, but very little livestock. Other households may have a short travel time but make frequent trips to their source. Because of this, I considered drinking water travel time and drinking water source a more accurate and consistent way to characterize the Water major component.

4.E. Comparison Between Groups

After following all of the previous steps for the assignment of a unique LVI score for each household, Welch's t-test, analysis of variance (ANOVA), and Tukey Honest Significant Difference (HSD) tests were applied between different groups. The groupings utilized head of household and country, and are as follows:

- Welch's t-test between overall data from MHH and FHH
- ANOVA and Tukey HSD test between all FDJ, FDF, MSP, and MNS households
- Welch's t-test between MHH and FHH within country subsets of data.
- ANOVA and Tukey HSD tests between FDJ, FDF, MSP, and MNS households within country subsets of data.

As stated in the Literature Review (*pg. 14*), FDJ households refer to those that are not associated with a male partner, due to the female head being widowed, divorced or never married. FDF households refer to those associated with a male partner, who may still be providing support of some kind to the household (Rogan, 2013). Welch's t-tests were performed using *Eq. (12)*:

(12)

$$t = \frac{\underline{x}_M - \underline{x}_W}{\sqrt{\frac{s_M^W}{N_M} + \frac{s_W^W}{N_W}}}$$

Where M represents MHH, W represents FHH, $\underline{x}$ is the sample mean, s the sample standard deviation and N the sample size. Analysis of variance (ANOVA) tests were performed to test the following null hypothesis listed in *Eq. (12)*:

$$H_0: \mu_{FDJ} = \mu_{FDF} = \mu_{MSP} = \mu_{MNS} \tag{13}$$

Where μ is the group mean for Female de jure (FDJ), Female de facto (FDF), Male with spouse MSP, and Male without spouse (MNS) households. The alternative hypothesis (H_A) that two (or more) group means are significantly different from the one another was accepted if the ANOVA returned a statistically significant result. Additionally, all ANOVA results were interpreted with consideration of the F-statistic. Tukey HSD tests were performed using *Eq. (14):*

$$HSD = \frac{M_x - M_y}{\sqrt{\frac{MS_w}{n_{HH}}}} \tag{14}$$

Where HSD represents the honest significant difference, $M_x - M_y$ is the pair of means being evaluated. MS_w is the mean square within, and n_{HH} is the number of households, which in this case is four.

4.F. Supporting analysis

In addition to calculated LVI and major component scores, other variables were considered to support speculations of conservancy impact on the studied communities. These variables include a gendered breakdown of employment, institutional aid, and conservancy participation.

4.F.i. Employment

I calculated a gendered breakdown of employment to better characterize how community-related employment may be benefiting different groups. The presented survey asked the occupation title of all studied household residents. Twenty-two broad job titles were given, these were categorized into four bins, Conservancy and Government Jobs, Public Jobs, Traditional Jobs, and Unknown (*Table 7, pg. 39*).

Conservancy and Government jobs are defined as employment directly provided by these entities. Public jobs are defined as those connected to public spaces of business. These jobs are noted as important in that they have a recognizable opportunity to increase directly with conservation activities, such seen through ecotourism (Khatiwada & Silva, 2015). Traditional jobs are defined by those that are connected to rural livelihoods in the study area and see less direct impact from the conservancy. After classification, I calculated gendered proportions of participants in each job category using *Eq. (15)*:

$$JP_W = \frac{J_w}{J_w + J_M} X\ 100 \tag{15}$$

Where JP represents the job percentage held by the assessed group, J_w is jobs held by women and J_M is jobs held by men. I also broke down job holdings into percentages by gender of household head using *Eq. (16)*.

$$JP_{FDJ} = \frac{J_{FDJ}}{J_{FDJ} + J_{FDF} + J_{MSP} + J_{MNS}} X\ 100 \tag{16}$$

Where JP represents the job percentage held by the assessed group.

Table 7: Gendered Employment Assessment			
Job Category	***Job Title***	***n Female***	***n Male***
Conservancy/ Government Job	Safari Guide	0	3
	National Service	22	5
	Escort Guide/ Conservancy Representative	3	6
	Government Technical Officer	9	5
Public Job	Driver	1	7
	Small Business Owner	62	22
	Lodge Employee	12	11
	Clerk (shop, office, bank)	12	7
	Chef	3	0
	Cleaner	1	10
	Security Guard	3	13
	Mechanic	0	1
	Manager	1	2
Traditional Job	Herder	6	32
	Teacher	16	13
	Casual Laborer	5	6
	Builder	1	6
	Nurse/Health Assistant	4	7
	Police	3	1
	Headman	0	7
	Soldier	0	0
	Miner	0	0
Unknown	Other	3	14

4.F.ii. Institutional Aid

I conducted a gendered analysis of institutional aid in order to identify groups that may be targeted for support. To do so, I assessed differences in money received as aid from government and CBC entities through an ANOVA test [*Eq. (13), pg. 37*]. Groups were defined as FDJ, FDF, MSP, and MNS households.

4.F.iii. Conservancy Participation

I assessed conservancy participation to identify any differences in participation by gender and to characterize the relationship between community participation and LVI scores. Community participation was defined through the survey with three questions, which asked if the household has heard of the conservancy, if the household attends conservancy meetings, and if the household is a member of the conservancy. These binary answers were used directly in analysis, and through combining answers to generate a cumulative score per household. This score consisted of a sum of each answer, 0 for no, 1 for yes to having heard of conservancy, 2 for yes to attend meetings, and 3 for yes to member of the conservancy.

To characterize conservancy participation, I found gendered percentages of households that have heard of the conservancy, attend conservancy meetings, and are members of the conservancy. In addition, I tested for significant differences between FHH and MHH in gendered participation scores utilizing a t.test [*Eq. (12), pg. 37*]

I tested for relationships between community participation and LVI overall and major component scores utilizing Pearson's product-moment correlation test, demonstrated in *Eq. (17):*

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2] - [n\Sigma y^2 - (\Sigma y)^2]}} \tag{17}$$

Where r indicates the correlation of the two variables, x indicates community participation, and y indicates the other component being tested for correlation. Each major component and the overall LVI score were tested to see if it demonstrated a correlation with the cumulative household conservancy participation score.

5. Results

5.A. Overall LVI Scores

While no significant differences were found in overall LVI scores between FHH and MHH, significant differences were found in the analysis of major components and further by household head status classifications. *Figures 4* and *5 (pg. 48)* demonstrate the gendered breakdown of the overall LVI average score in the major components, and *Figure 6* (*pg. 48*) demonstrates the breakdown of this score by country. P-values from the t-test and Tukey HSD assessment are listed in *Table 8 (pg. 48)*. All major components are discussed below in their respective section.

5.B. Breakdown of Major Component Results by Household

5.B.i. Food

T-tests of the overall dataset indicated that FHH are significantly more food insecure when compared to MHH. Tukey HSD household head groupings revealed that FDJ households experience the most food insecurity when compared to FDF and MSP households (*Figure 10, pg. 49)*. However, vulnerabilities were also found in MNS households when compared to FDF and MSP households.

5.B.ii. Wealth

T-tests of the overall dataset did not demonstrate any significant differences in wealth between FHH and MHH. These results were repeated in the ANOVA and Tukey HSD tests, with no household category showing significant vulnerabilities when compared to other categories.

5.B.iii. Socio-Demographic

T-tests of the overall dataset indicated that FHH demonstrates more socio-demographic related vulnerabilities when compared to MHH. Tukey HSD household head groupings revealed that differences are found primarily between FDJ and MSP households (*Figure 11, pg. 49*).

5.B.iv. Health

T-tests of the overall dataset did not demonstrate any significant differences in health between FHH and MHH. These results were repeated in the ANOVA and Tukey HSD tests, with no household category showing significant vulnerabilities when compared to other categories.

5.B.v. Social Networks

T-tests of the overall dataset indicated that FHH demonstrates more social network related vulnerabilities when compared to MHH. Tukey HSD household head groupings revealed that differences are found primarily between FDJ and MSP households (*Figure 12, pg. 49*).

5.B.vi. Water

T-tests of the overall dataset indicated that MHH demonstrates water related vulnerabilities when compared to FHH. Tukey HSD household head groupings revealed that differences are found primarily between more vulnerable MSP and FDF households when compared to less vulnerable FDJ households.

5.B.vii. Livelihood Diversification

T-tests of the overall dataset indicated that MHH demonstrates less indicators of livelihood diversification when compared to FHH. Tukey HSD household head groupings revealed that differences are found primarily between more vulnerable MSP households when compared to less vulnerable FDJ and MNS households (*Figure 13, pg. 49*)

5.B.viii. Environmental Shock

T-tests of the overall dataset did not demonstrate any significant differences in health between FHH and MHH. These results were repeated in the ANOVA and Tukey HSD tests, with no household category showing significant vulnerabilities when compared to other categories.

5.C. Breakdown of Major Component Results by Household

5.C.i. Botswana

Data from Botswana demonstrated overall higher LVI scores for FHH in that country (*Appendix B, pg. 73)*. However, Tukey HSD tested [*Eq. (13), pg. 37*] household groupings did not reveal significant differences between any particular household groups. In addition, Botswana mirrored patterns for increased vulnerabilities of FDJ households in food and socio demographic components when compared to MSP households. Botswana did not demonstrate any further significant differences (*Figure 7, pg. 48*).

5.C.ii. Namibia

Data from Namibia demonstrated overall higher LVI scores for FHH in that country. Tukey HSD tested [*Eq. (13), pg. 37*] household groupings revealed that differences are primarily found between more vulnerable FDJ households when compared to less vulnerable FDF and MSP households (*Appendix B, pg. 73)*. In addition, Namibia demonstrated increased vulnerability of FDJ and FDF households in the food major component when compared to MNS households. Other notable differences include FDJ vulnerabilities in the socio demographic component when compared to FDF households and MNS vulnerabilities in health components when compared to FDJ households (*Figure 8, pg. 48*)

5.C.iii. Zambia

Data from Zambia demonstrated no significant differences in overall LVI scores between FHH and MHH (*Appendix B, pg. 73*). Tukey HSD tested [*Eq. (13), pg. 37*] household groupings revealed increasing vulnerabilities of FDJ households in the socio demographic component when compared to FDF and MSP households, as well as FDF vulnerabilities when compared to MSP households. Other notable differences include increased environmental shock vulnerabilities of MSP vulnerabilities when compared to FDJ households (*Figure 9, pg. 48*).

5.D. Supporting Analysis

5.D.i. Employment

The highest percentage of jobs reported are held by women (51%), followed closely by men (49%). The majority of jobholders are from FDJ households (44%), followed by MSP (39%), then MNS (11%), and lastly FDF (6%) (*Figure 14, pg. 51*).

Additionally, women hold the highest percentage of Conservancy and Government jobs (64%) and Public jobs (62%), while men hold the highest percentage of Traditional jobs (67%). The majority of Conservancy and Government employees are from FDJ households (49%), as well as the majority of Public jobs (58 %) (*Table 7, pg. 39).* Comparatively, the highest number of traditional jobs are from MSP households (46%). *Figures 15* and *16* (*pg. 51)* visually demonstrate the breakdown of employment.

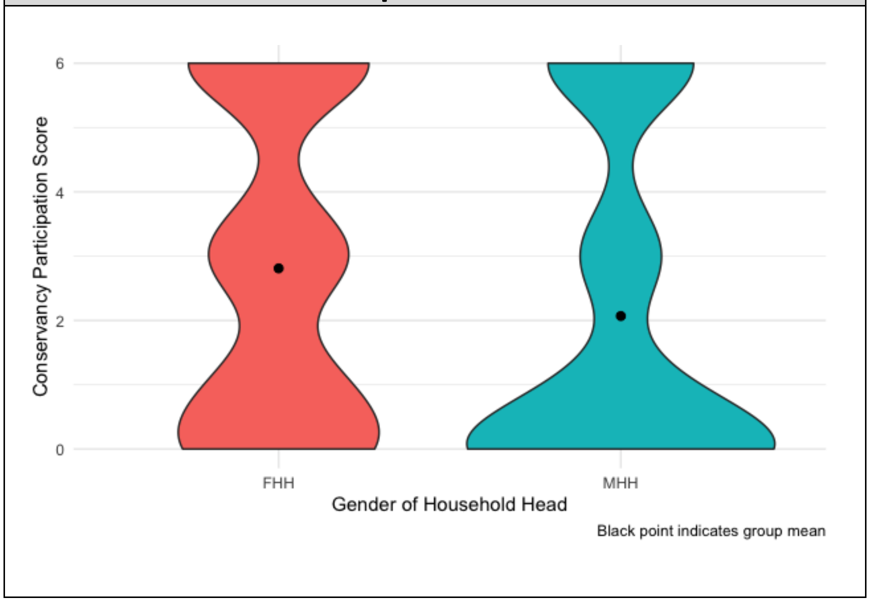
5.D.ii. Institutional Aid

No significant differences were found in governmental, conservancy, or combined aid between household groups.

5.D.iii. Conservancy Participation

I assessed gendered conservancy participation by finding gendered percentages of households that have heard of the conservancy, attend conservancy meetings, and are members of the conservancy. The majority of households who have heard of the conservancy were female headed (56%) (*Figure 17*, pg. 51). Likewise, the majority of households attending meetings were also female headed (55%) (*Figure 18, pg. 51)*. However, MHH comprised the majority of conservancy members (78%) (*Figure 19, pg. 15*). In addition, I tested for significant differences between FHH and MHH answers to these questions and to gendered participation scores utilizing a t-test [*Eq. (11), pg. 34*]. FHH were found to have significantly higher conservancy participation scores than MHH (*Figure 3*).

Figure 3: Gendered Conservancy Participation Scores



Results of a correlation tests ran between conservancy participation score and major components yielded mixed results (*Table 9, pg. 48).* Negative correlations were found between conservancy participation scores and the following major components: health, livelihood diversification, and water. This indicates that higher conservancy participation scores are associated with lower vulnerability in these categories. Positive correlations were found between community participation scores and major components food and social network. This indicates that higher conservancy participation scores are associated with lower vulnerability in these categories. No correlation was found between conservancy participation and overall LVI scores or the following major components; household wealth, environmental shock, and socio-demographic.

Table 8: Significant Differences in Households
Listed as P-values from three different statistical tests

	Groups	LVI	FC	HWC	SDC	HC	SNC	LDC	WC	ESC
T-test	MHH – FHH	0.619	0.0001984 ***FHH	0.35396	0.001577 ***FHH	0.7133	0.01339 ***FHH	3.792e-05 ***MHH	6.461e-07 ***MHH	0.3454
ANOVA[1]	All	0.841	5.59e-07 ***	0.887	0.0154 ***	0.327	0.0358 ***	2.42e-08 ***	5.07e-06 ***	0.186
Tukey HSD	FDJ-FDF	0.9931629	0.0435004 ***FDJ	0.9629338	0.9381442	0.2656755	0.8298736	0.7408495	0.0239970 ***FDF	0.9673716
	MNS-FDF	0.9999708	0.0452213 ***MNS	0.9662861	0.8616556	0.6530846	0.9425600	0.6979380	0.3719684	0.9702869
	MSP-FDF	0.9523129	0.9683627	0.8941636	0.8833294	0.4306749	0.9805329	0.4100938	0.8956427	0.7371739
	MNS-FDJ	0.9788552	0.9105869	0.9996174	0.2351817	0.9567471	0.9969082	0.9824449	0.6955280	0.6044987
	MSP-FDJ	0.9331955	0.0000038 ***FDJ	0.9549531	0.0131468 ***FDJ	0.8916828	0.0225692 ***FDJ	0.0000000 ***MSP	0.0000056 ***MSP	0.5689957
	MSP-MNS	0.8907573	0.0047248 ***MNS	0.9981381	0.9932248	0.9992116	0.5501514	0.0024490 ***MSP	0.4100567	0.2172236

***** Indicates significance followed by the abbreviation for the more vulnerable group*

[1] *All ANOVA results interpreted with consideration to the F-statistic*

Table 9: Results of Pearson's Product-Moment Correlation Test

PCN &:	LVI	FC	HWC	SDC	HC	SNC	LDC	WC	ESC
P-value	0.4239	1.956e-07 ***	0.2657	0.9955	0.000106 ***	3.543e-09 ***	5.066e-14 ***	0.001878 ***	0.1692
Correlation	-0.029725	0.191656	-0.0413658	-0.0002083	-0.143366	0.216887	-0.274558	-0.115198	-0.0510808

**** Indicates significance*

Figure 4: LVI by Gender of Household Head	Figure 5: LVI by Gendered Household Groups	Figure 6: LVI By Country

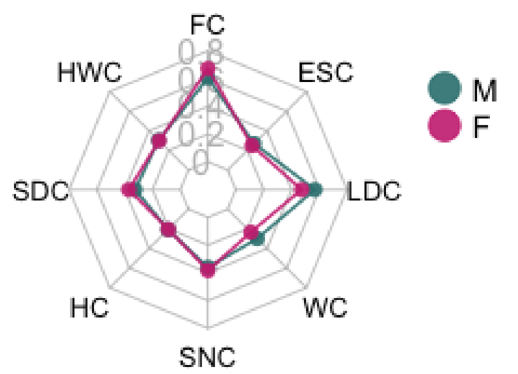
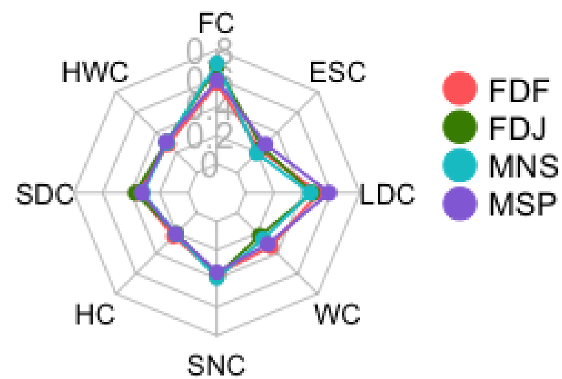
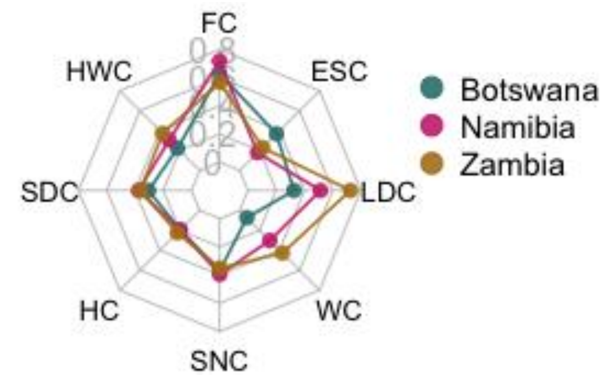

		

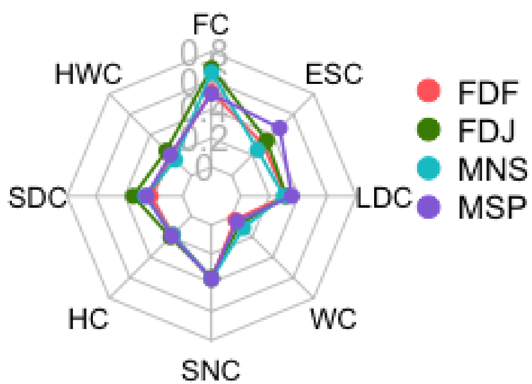
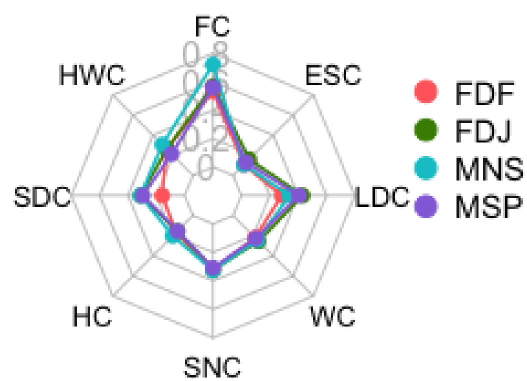
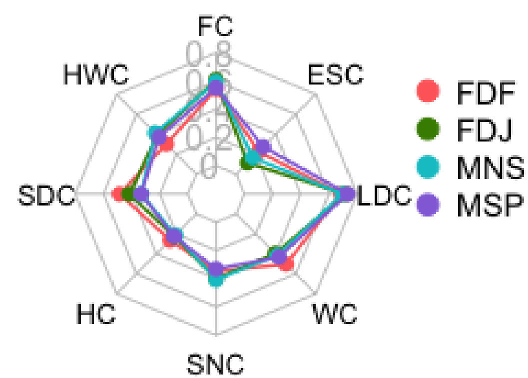
Figure 7: Botswana LVI by Gendered Household Groups	**Figure 8: Namibia LVI by Gendered Household Groups**	**Figure 9: Zambia LVI by Gendered Household Groups**
		

Figure 10: Food Component by Household

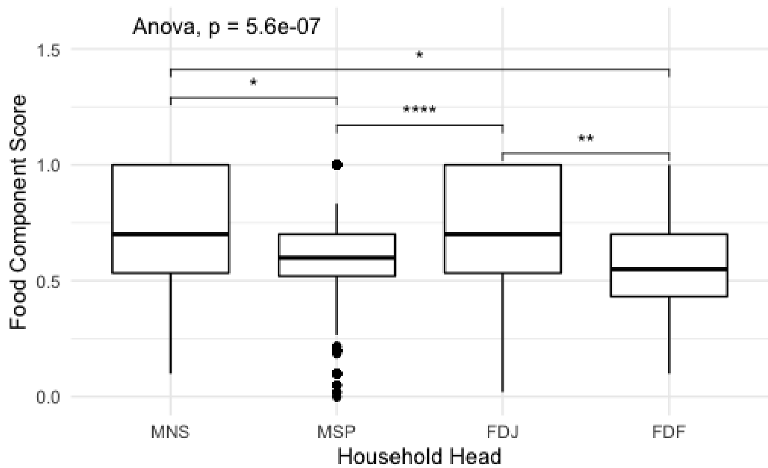


Figure 11: Socio-Demographic Component by Household

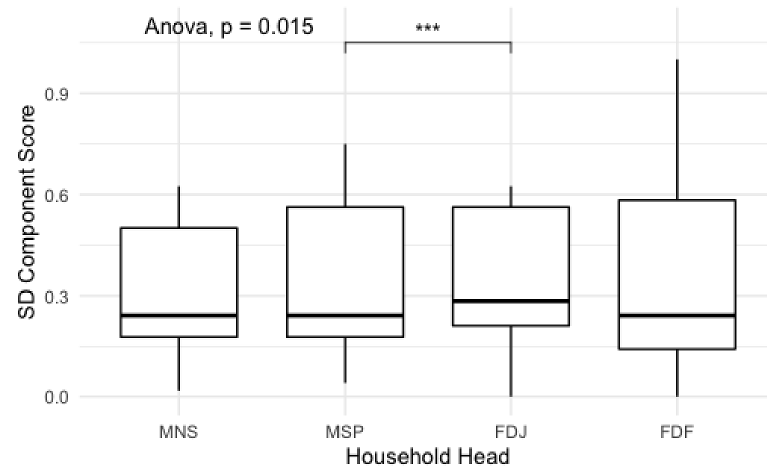


Figure 12: Social Network Component by Household

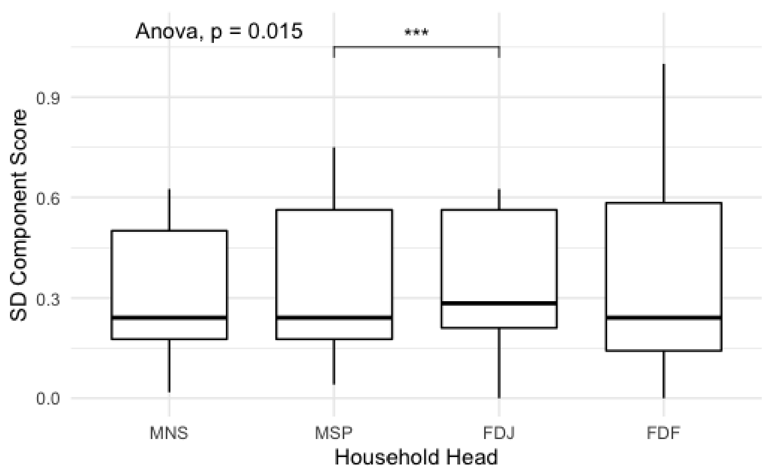


Figure 13: Livelihood Diversification Component by Household

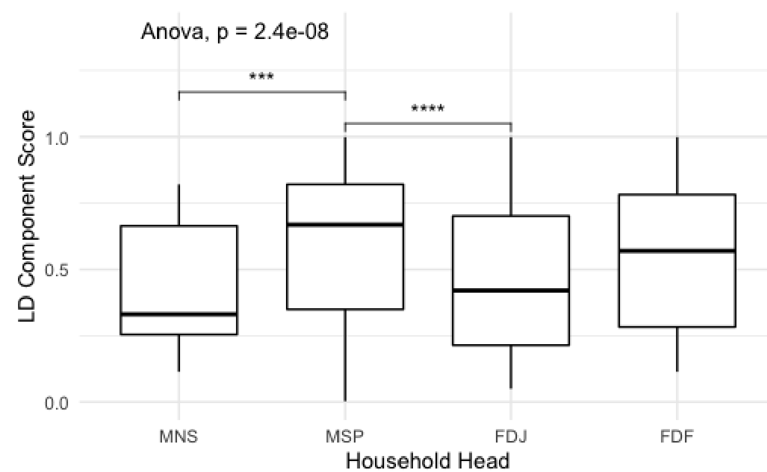


Figure 14: Employment by Household Type

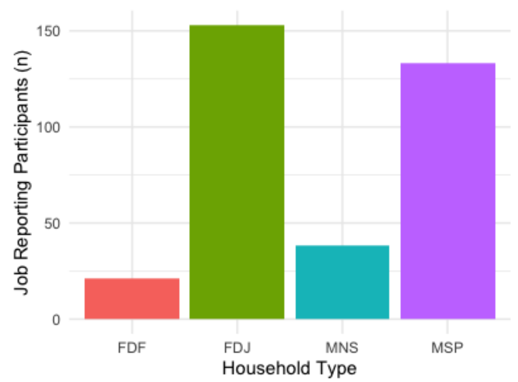


Figure 15: Employment Breakdown by Gender of Job Holder

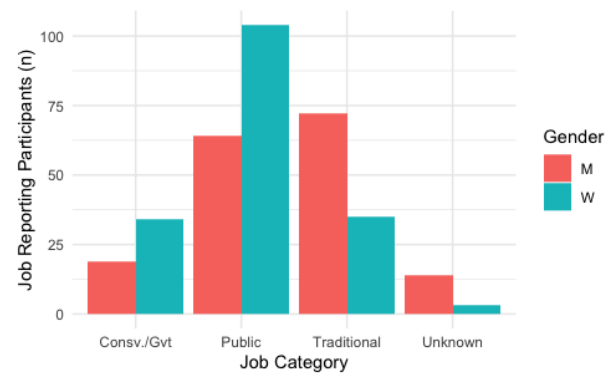


Figure 16: Employment Breakdown by Household Type

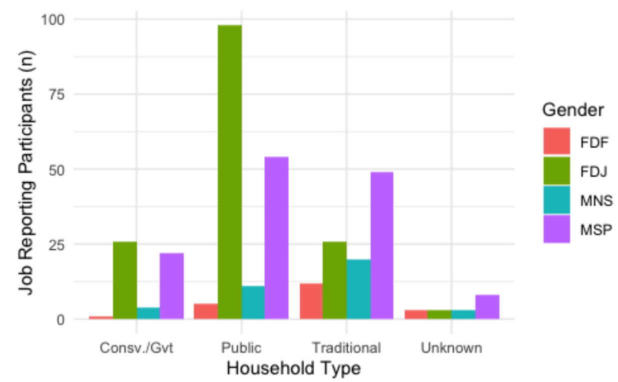


Figure 17: Household "Heard of Conservancy" Response

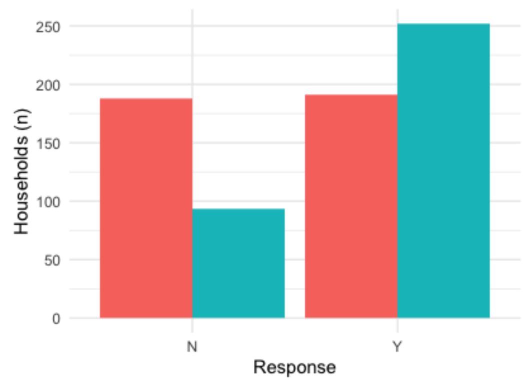


Figure 18: Household "Attend Conservancy Meetings" Response

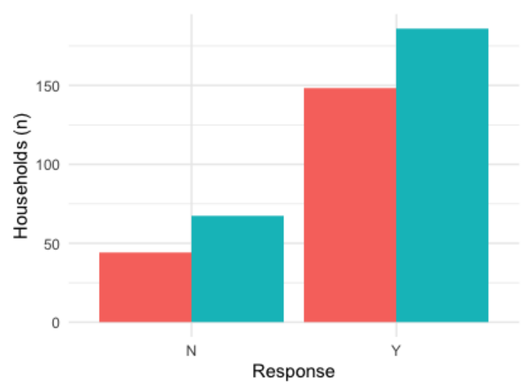
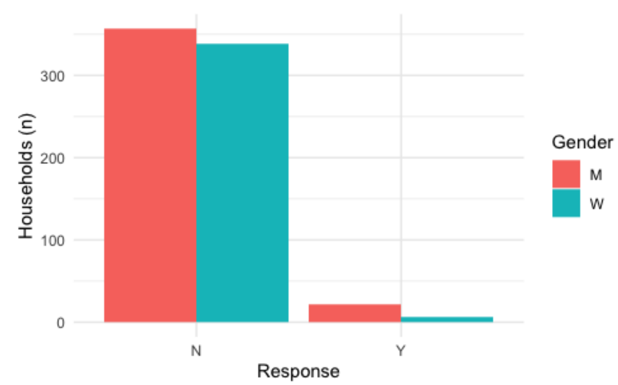


Figure 19: Household "Member of Conservancy" Response



6. Discussion

The results of this study demonstrate that more indicators of gender equality in regard to climate vulnerability exist within our focal group when compared to those in previous studies. Although the direct impact of the conservancy is unclear, it is likely that it has influenced these trends through increased employment for women and FDJ households. In the following paragraphs, I will further explore disparities in vulnerabilities found and provide supporting information of how the unique study setting in a CBC may be impacting and altering expected data trends.

Literature has indicated that FHH of SSA are particularly vulnerable to climate change when compared to male headed counterparts (*Table 1, pg. 16*). This can be said to be particularly true of FDJ households, who are not associated with a male partner, unlike FDF households (Rogan 2013). Previous studies have found increased vulnerabilities in FHH that relate to each of the chosen major components (*Table 2, pg. 19*). This study demonstrates that a few of these previously observed patterns exist, particularly in terms of vulnerabilities associated with social networks, socio-demographics, and food insecurity in FDJ headed households. Lower socio-demographic scores are supported by previous studies asserting that that female household heads generally have lower education and households that include more elderly and children members (Tibesigwa & Visser, 2016; Dungumaro, 2008; Rogan, 2013; Alhassan et al., 2019). Lower social-network scores are supported by previous studies asserting that heads of households are often restricted from community decision making and have limited access to credit (Alhassan et al., 2019; Balehey et al., 2018). Finally, trends of higher food insecurity is well supported in literature as an issue that disproportionately impacts FHH in

SSA (Tibesigwa & Visser, 2016; Dungumaro, 2008; Alhassan et al., 2019; Rao et al., 2020). As these trends persist predominately in FDJ households, this group continues to necessitate the most attention from conservancy and governmental institutions. Currently, there are no significant differences in governmental and conservancy aid received by this group when compared to other household types. This identifies an area for improvement, as funding targeted at FDJ households may help alleviate some of the recognized disparities, such as food aid to help alleviate amplified food insecurity.

Other expected patterns of gendered disparities do not exist in this study group. No significant differences were found in overall LVI scores or household wealth, health, and environmental shock component scores between FHH and MHH. In addition, MHH unexpectedly displayed increased vulnerabilities related to water and livelihood diversification. Employment data for survey respondents may help to provide insight into why many gendered disparities do not seem to be present within the surveyed community. An employment breakdown showed that female participants held the greatest number of jobs at 51%, and that the majority of total reported jobs were also held by members of FDJ households at 44%. This gendered breakdown is higher than country reporting's for the female percentage of total workforce, which is 48.4% for Botswana, 48.6% for Namibia, and 48.3% for Zambia (World Bank, 2019). Of the jobs held by survey participants, 15% were conservancy and government provided, 49% were public jobs with potential connection to the conservancy, 31% were traditional, and 5% were unknown. Female participants and FDJ households seem to benefit most from the presence of conservancy and government employment, holding the highest majority of these jobs. In addition, female participants and FDJ households hold the majority of public jobs. This is significant because of previous observations that show public job opportunities increase with conservancy and ecotourism activity (Khumalo & Freimund, 2014; Khatiwada & Silva, 2015). These job titles include those related to increased infrastructure for tourists, such as clerks, lodge employees, cleaners, chefs, and security guards, or as small business owners to sell goods to incoming community visitors. As expected, traditional jobs were held predominately by male participants (67%) and by MSP households (46%).

This summary of employment is especially meaningful when considering household wealth and livelihood diversification. As noted, no significant differences were found in the Household Wealth component. This contradicts expected data trends, as lower wealth in FHH relates to the “triple burden” of female labor market disadvantages, time constraints of balancing household tasks with employment, and the lack of a partner who is also earning wages (Rogan, 2013). Increased employment by conservancies for female participants may be influencing increased income and material wealth. In addition, increased employment opportunities may have led to significantly higher livelihood diversification for FHH compared to male counterparts. Livelihood diversification, or the engagement in multiple economic activities, reduces risk related to loss of income and subsequently improves household security. Livelihood diversity is often addressed in literature as an important buffer against climate vulnerability due to its positive impact on adaptive capacity (Adzawla et al., 2019; Rao et al., 2020). Livelihood diversification has also been linked to increases in food security, a trend that could potentially appear in the study set over time (Adzawla et al., 2019). Previously conducted studies support these findings and report that while wealth attributed to increased livelihood diversification and employment opportunities often increases for FHH in conservation areas, there are no significant changes in wealth for MHH (Khatiwada & Silva, 2015; Khumalo & Freimund, 2014). Higher adaptive capacity to climate related shocks is directly attributed to increased livelihood diversification and wealth. (Rao et al., 2020; Adzawla et al., 2019; Ncube et al., 2016; Alhassan et al., 2019; Balehey et al., 2018).

Indicators that the conservancy may be impacting lessened vulnerability in female groups may also be found in analyses of conservancy participation. Results found FHH to be significantly more involved in the conservancy. Higher participation scores were found to correlate with greater health, livelihood diversification, and water security. Although correlation does not signify causation, it is interesting to note that the presence of the conservancy may be related to heightened FHH scores in these categories.

7. Conclusion

In conclusion, a growing body of literature supports that FHH and female individuals are disproportionately vulnerable to the shocks of climate change in SSA (*Table 1, pg. 16).* This development in gender inequality is critical to consider in the face of severe and increasing climate shocks projected to substantially impact SSA (Serdeczny et al., 2017; Alagidede et al., 2016; Brooks, 2018; Balehey et al., 2018; Connolly-Boutin & Smit, 2016).

To test whether these observed patterns exist within the study area in a unique CBC-setting, I assembled an LVI comprised of eight major components. This system of measurement, initially developed by (Hahn et al., 2009) attempts to fully encapsulate the many dimensions of climate vulnerability, including food insecurity, household wealth, socio-demographic profile, social network accessibility, livelihood diversification, and environmental shocks. Index scores of FHH and MHH were tested for significant differences using Welch's t-test, ANOVA, and Tukey HSD tests.

My results found that many expected trends persist, while others did not. FDJ households exhibited the most vulnerabilities overall, namely in food insecurity, socio-demographics, and social networks. However, no significant differences were found between overall MHH and FHH regarding household wealth, environmental shock, and health. Furthermore, MHH unexpectedly demonstrated more vulnerability in regard to livelihood diversification, suggesting that FHH have increased systems of supporting their households in comparison to MHH. These findings are significant, as household adaptive capacity is attributed largely to wealth and diversification of livelihood (Rao et al., 2020; Adzawla et al., 2019; Ncube et al., 2016; Alhassan et al., 2019; Balehey et al., 2018). In turn, increased household adaptive capacity indicates an improved aptitude of households to withstand complete destruction of material wealth, loss of income, and degradation of human health during climate-related shocks (Turner et al., 2003; Ncube et al., 2016).

Based on my analysis, I suggest that CBC's are making an impactful difference on the livelihood diversification and wealth of FHH. This is due to employment opportunities directly provided by and indirectly created by the CBC. Most direct conservancy and government jobs are held by female employees and members of FDJ households. Furthermore, this impact of CBC's has been found to positively impact wealth and livelihood diversification studies in Namibia (Khatiwada & Silva, 2015; Khumalo & Freimund, 2014). Finally, other supporting analysis conducted in this study found a positive correlation between livelihood diversification and a compiled conservancy participation score.

The results of this study reveal many areas that could be expanded upon with future research. Both a strength and a weakness of this study is its data-driven nature. This eliminates researcher bias, but also prevents deeper contextual knowledge of the local area and traditions. Suggestions for future studies include incorporating qualitative information from focus groups on local customs, gender inequality, and gendered climate vulnerabilities within the focal communities. Additionally, previous studies served as a stand in for a control group in my research. I propose implementation of the LVI framework on other areas of Namibia, Botswana, and Zambia outside of KAZA to fully characterize the impact of the CBC on gender. Finally, although regional differences were present in the data analysis, such results were not a focus of the study. Data from the country subsets could be scrutinized further in depth with more contextual information about each unique country.

As stated before, this study aims to deepen understanding of the relationship between gendered vulnerabilities and climate, to identify vulnerable populations within our studied communities, and to provide insight on how policy and employment opportunity can be used to promote climate justice in rural areas. My hope for this research is that it will continue to add to knowledge about how CBC's are found to address the complex issues associated with balancing conservation and human livelihoods, while promoting equality and opportunity for overlooked groups.

Acknowledgements

I would like to extend a huge thank you to the people of our focal communities for their participation in the study, as well as local authorities and participants who helped inform the data-collection process.

Furthermore, I would like to thank the KAZAVA team for their tireless efforts in collected data through conducting household surveys, and the local enumerators who provided translating services.

In particular, I would like to thank my mentor Michael Drake (Ph.D. student at CU Boulder in Environmental Studies) of the KAZAVA team, who offered me the opportunity to work with the dataset and spent an enormous amount of time guiding my research and grant writing, assisting me in refining my ideas, and troubleshooting technical issues.

A big thank you to Sarah Kelly, my thesis advisor, who provided support through editing, statistical advising, assistance writing my thesis proposal, and scheduling deadlines for various drafts.

Thank you also to Joel Hartter, who is both a committee member and a member of the KAZAVA team. Dr. Hartter provided guidance and assistance throughout the process of obtaining a UROP grant to fund the preliminary stages of this project. His involvement was also crucial in opening up the opportunity to join KAZAVA team and work with the data.

Other honorable mentions of the KAZAVA team include Karen Bailey, who took time to evaluate my project at various stages and give thoughtful and meaningful critiques and suggestions, Jon Salerno, who provided the R-code for the food insecurity portion of my project, and Lin Cassidy who provided resources to help inform my methods of characterizing material wealth within the study area.

Additionally, I would like to thank William Travis for agreeing to be my honors council representative and answering many questions about the thesis process.

Finally, gratitude to all others who helped support me in my writing and research process, including family, friends, and loved ones

References

Adzawla, W., Baumüller, H., Donkoh, S. A., & Serra, R. (2019). Effects of climate change and livelihood diversification on the gendered productivity gap in Northern Ghana. *Climate and Development*, *0*(0), 1–13. https://doi.org/10.1080/17565529.2019.1689093

African Financial Markets Initiative (AFMI) (2020). *Southern Africa.* Retrieved from https://www.africanbondmarkets.org/en/country-profiles/southern-africa/

Aheto, D. W., Kankam, S., Okyere, I., Mensah, E., Osman, A., Jonah, F. E., & Mensah, J. C. (2016). Community-based mangrove forest management: Implications for local livelihoods and coastal resource conservation along the Volta estuary catchment area of Ghana. *Ocean & Coastal Management*, *127*, 43–54. https://doi.org/10.1016/j.ocecoaman.2016.04.006

Alagidede, P., Adu, G., & Frimpong, P. B. (2016). The effect of climate change on economic growth: Evidence from Sub-Saharan Africa. *Environmental Economics and Policy Studies*, *18*(3), 417–436. https://doi.org/10.1007/s10018-015-0116-3

Alhassan, S. I., Kuwornu, J. K. M., & Osei-Asare, Y. B. (2019). Gender dimension of vulnerability to climate change and variability: Empirical evidence of smallholder farming households in Ghana. *International Journal of Climate Change Strategies and Management*, *11*(2), 195–214. https://doi.org/10.1108/IJCCSM-10-2016-0156

Assan, E., Suvedi, M., Olabisi, L. S., & Allen, A. (2018). Coping with and Adapting to Climate Change: A Gender Perspective from Smallholder Farming in Ghana. *Environments*, *5*(8), 86. https://doi.org/10.3390/environments5080086

Baffoe, G., & Matsuda, H. (2018). An Empirical Assessment of Households Livelihood Vulnerability: The Case of Rural Ghana. *Social Indicators Research*, *140*(3), 1225–1257. https://doi.org/10.1007/s11205-017-1796-9

Bajracharya, S. B., Furley, P. A., & Newton, A. C. (2006). Impacts of Community-based Conservation on Local Communities in the Annapurna Conservation Area, Nepal. *Biodiversity & Conservation; Dordrecht*, *15*(8), 2765–2786. http://dx.doi.org.colorado.idm.oclc.org/10.1007/s10531-005-1343-x

Balehey, S., Tesfay, G., & Balehegn, M. (2018). Traditional gender inequalities limit pastoral women's opportunities for adaptation to climate change: Evidence from the Afar pastoralists of Ethiopia. *Pastoralism-Research Policy and Practice*, *8*, 23. https://doi.org/10.1186/s13570-018-0129-1

Balikoowa, K., Nabanoga, G., & Tumusiime, D. M. (2018). Gender stereotyping: Evidence from gender differentiated household vulnerability to climate change in Eastern Uganda. *Cogent Environmental Science*, *4*(1), 1512838. https://doi.org/10.1080/23311843.2018.1512838

Berkes, F. (2007). Community-based conservation in a globalized world. *Proceedings of the National Academy of Sciences*, *104*(39), 15188–15193. https://doi.org/10.1073/pnas.0702098104

Blaikie, P. (2006). Is Small Really Beautiful? Community-based Natural Resource Management in Malawi and Botswana. *World Development*, *34*(11), 1942–1957. https://doi.org/10.1016/j.worlddev.2005.11.023

Born, L., Spillane, C., & Murray, U. (2019). Integrating gender into index-based agricultural insurance: A focus on South Africa. *Development in Practice*, *29*(4), 409–423. https://doi.org/10.1080/09614524.2018.1556608

Brooks, K. (2018). Rising absolutely, declining relatively: Agriculture, climate change, and job creation in Africa. *Agrekon*, *57*(3–4), 181–197. https://doi.org/10.1080/03031853.2018.1538002

Connolly-Boutin, L., & Smit, B. (2016). Climate change, food security, and livelihoods in sub-Saharan Africa. *Regional Environmental Change*, *16*(2), 385–399. https://doi.org/10.1007/s10113-015-0761-x

Demographic and Health Surveys Program (DHS Program), (2020). *Wealth Index Construction.* United States Agency of International Development (USAID). Retrieved from https://www.dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm.

Deressa, T., Hassan, R. M., Ringler, C., Alemu, T., & Yesuf, M. (2009). Determinants of farmers' choice of adaptation methods to climate change in the Nile Basin of Ethiopia. *Global Environmental Change*, *19*(2), 248–255. https://doi.org/10.1016/j.gloenvcha.2009.01.002

Dungumaro, E. W. (2008). Gender Differentials in Household Structure and Socioeconomic Characteristics in South Africa. *Journal of Comparative Family Studies*, *39*(4), 429-+. https://doi.org/10.3138/jcfs.39.4.429

Fauchereau, N., Trzaska, S., Rouault, M., & Richard, Y. (2003). Rainfall variability and changes in Southern Africa during the 20th century in the global warming context. *Natural Hazards*, *29*(2), 139–154. https://doi.org/10.1023/A:1023630924100

Flato, M., Muttarak, R., & Pelser, A. (2017). Women, Weather, and Woes: The Triangular Dynamics of Female-Headed Households, Economic Vulnerability, and Climate Variability in South Africa. *World Development*, *90*, 41–62. https://doi.org/10.1016/j.worlddev.2016.08.015

Fuller, R., & Lain, J. (2019). Are female-headed households less resilient? Evidence from Oxfam's impact evaluations. *Climate and Development*. https://doi.org/10.1080/17565529.2019.1637330

Gender Affairs Department (GAD), (2014). *Botswana Country Report on the Implementation of the Beijing Platform for Action (Beijing Plus 20 Years).* Retrieved from https://sustainabledevelopment.un.org/content/documents/13016Botswana_review_Beijing20.pdf.

Goldman, M. J., & Little, J. S. (2015). Innovative Grassroots NGOS and the Complex Processes of Women's Empowerment: An Empirical Investigation from Northern Tanzania. *World Development*, *66*, 762–777. https://doi.org/10.1016/j.worlddev.2014.09.005

Hackel, J. D. (1999). Community Conservation and the Future of Africa's Wildlife. *Conservation Biology*, *13*(4), 726–734. https://doi.org/10.1046/j.1523-1739.1999.98210.x

Hahn, M. B., Riederer, A. M., & Foster, S. O. (2009). The Livelihood Vulnerability Index: A pragmatic approach to assessing risks from climate variability and change—A case study in Mozambique. *Global Environmental Change*, *19*(1), 74–88. https://doi.org/10.1016/j.gloenvcha.2008.11.002

Hoscilo, A., Balzter, H., Bartholome, E., Boschetti, M., Brivio, P. A., Brink, A., Clerici, M., & Pekel, J. F. (2015). A conceptual model for assessing rainfall and vegetation trends in sub-Saharan

Africa from satellite data. *International Journal of Climatology*, *35*(12), 3582–3592. https://doi.org/10.1002/joc.4231

Ibrahim, S. S., Ozdeser, H., & Cavusoglu, B. (2019). Vulnerability to recurrent shocks and disparities in gendered livelihood diversification in remote areas of Nigeria. *Environmental Science and Pollution Research*, *26*(3), 2939–2949. https://doi.org/10.1007/s11356-018-3854-5

Internal Revenue Service (IRS), (2020). *Topic No. 400 Types of Income.* Retrieved from https://www.irs.gov/taxtopics/tc400.

Kakota, T., Nyariki, D., Mkwambisi, D., & Kogi-Makau, W. (2011). Gender vulnerability to climate variability and household food insecurity. *Climate and Development*, *3*(4), 298–309. https://doi.org/10.1080/17565529.2011.627419

Kavango Zambezi Transfrontier Conservation Area (KAZA TFCA), 2019. *Home*. https://www.kavangozambezi.org/en/.

Ketlhoilwe, M. J., & Kanene, K. M. (2018). Access to energy sources in the face of climate change: Challenges faced by women in rural communities. *Jamba-Journal of Disaster Risk Studies*, *10*, a375. https://doi.org/10.4102/jamba.v10i1.375

Khatiwada, L. K., & Silva, J. A. (2015). Mitigating Gender Inequality in Rural Regions: The Effects of Tourism Employment in Namibia. *International Journal of Tourism Research*, *17*(5), 442–450. https://doi.org/10.1002/jtr.2010

Khumalo, K. E., & Freimund, W. A. (2014). Expanding Women's Choices through Employment? Community-Based Natural Resource Management and Women's Empowerment in

Kwandu Conservancy, Namibia. *Society & Natural Resources*, *27*(10), 1024–1039. https://doi.org/10.1080/08941920.2014.905893

Kumssa, A., & Jones, J. F. (2010). Climate change and human security in Africa. *International Journal of Sustainable Development and World Ecology*, *17*(6), 453–461. https://doi.org/10.1080/13504509.2010.520453

Larson, L. R., Conway, A. L., Krafte, K. E., Hernandez, S. M., & Carroll, J. P. (2016). Community-based conservation as a potential source of conflict around a protected area in Sierra Leone. *Environmental Conservation*, *43*(3), 242–252. https://doi.org/10.1017/S0376892916000096

Lee, D. E., & Bond, M. L. (2018). Quantifying the ecological success of a community-based wildlife conservation area in Tanzania. *Journal of Mammalogy*, *99*(2), 459–464. https://doi.org/10.1093/jmammal/gyy014

Mason, S. J., & Jury, M. R. (1997). Climatic variability and change over southern Africa: A reflection on underlying processes. *Progress in Physical Geography*, *21*(1), 23–50. https://doi.org/10.1177/030913339702100103

Meinzen-Dick, R., Quisumbing, A., Doss, C., & Theis, S. (2019). Women’s land rights as a pathway to poverty reduction: Framework and review of available evidence. *Agricultural Systems*, *172*, 72–82. https://doi.org/10.1016/j.agsy.2017.10.009

Milazzo, A., & van de Walle, D. (2017). Women Left Behind? Poverty and Headship in Africa. *Demography*, *54*(3), 1119–1145. https://doi.org/10.1007/s13524-017-0561-7

Mugisha, A. R., & Jacobson, S. K. (2004). Threat reduction assessment of conventional and community-based conservation approaches to managing protected areas in Uganda.

Environmental Conservation, *31*(3), 233–241. https://doi.org/10.1017/S0376892904001432

Nanang, D. M., & Nunifu, T. K. (2010). *Natural Resources in Ghana: Management, Policy and EconomicsRUSH NEED BY APRIL RUSH 10 DAY BOOK*. Nova Science Publishers, Incorporated. http://ebookcentral.proquest.com/lib/ucb/detail.action?docID=3018798

Nchu, I. N., Kimengsi, J. N., & Kapp, G. (2019). Diagnosing Climate Adaptation Constraints in Rural Subsistence Farming Systems in Cameroon: Gender and Institutional Perspectives. *Sustainability*, *11*(14), 3767. https://doi.org/10.3390/su11143767

Ncube, M., Madubula, N., Ngwenya, H., Zinyengere, N., Zhou, L., Francis, J., Mthunzi, T., Olivier, C., & Madzivhandila, T. (2016). Climate change, household vulnerability and smart agriculture: The case of two South African provinces. *Jamba-Journal of Disaster Risk Studies*, *8*(2). https://doi.org/10.4102/jamba.v8i2.182

Nellemann, C., Verma, R., and Hislop, L., (2011). *Women at the frontline of climate change: Gender risks and hopes.* A Rapid Response Assessment. United Nations Environment Programme, GRID-Arendal.

Nwosu, C. O., & Ndinda, C. (2018). *Female household headship and poverty in South Africa: An employment-based analysis*. ERSA working paper 761.

Omolo, N., & Mafongoya, P. L. (2019). Gender, social capital and adaptive capacity to climate variability A case of pastoralists in arid and semi-arid regions in Kenya. *International Journal of Climate Change Strategies and Management*, *11*(5), 744–758. https://doi.org/10.1108/IJCCSM-01-2018-0009

Organisation for Economic Co-operation and Development (OECD) (2014). *Sub-Saharan Africa - The Social Institutions and Gender Index (SIGI) Regional Report*. https://www.oecd.org/dev/development-gender/Mini-Brochure_SIGI_AfriqueSubsaharienne_EN-web.pdf

Owusu, M., Nursey-Bray, M., & Rudd, D. (2019). Gendered perception and vulnerability to climate change in urban slum communities in Accra, Ghana. *Regional Environmental Change*, *19*(1), 13–25. https://doi.org/10.1007/s10113-018-1357-z

Pachauri, R. K., Mayer, L., & Intergovernmental Panel on Climate Change (Eds.). (2015). *Climate change 2014: Synthesis report*. Intergovernmental Panel on Climate Change.

Pallardy, R. (2012). Southern Africa's Kavango Zambezi Transfrontier Conservation Area. Retrieved from https://www.britannica.com/topic/Kavango-Zambezi-Transfrontier-Conservation-Area-1883803

Pons-Duran, C., Lucas, A., Narayan, A., Dabalen, A., & Menéndez, C. (2019). Inequalities in sub-Saharan African women's and girls' health opportunities and outcomes: Evidence from the Demographic and Health Surveys. *Journal of Global Health*, *9*(1). https://doi.org/10.7189/jogh.09.010410

Rao, N., Singh, C., Solomon, D., Camfield, L., Sidiki, R., Angula, M., Poonacha, P., Sidibé, A., & Lawson, E. T. (2020). Managing risk, changing aspirations and household dynamics: Implications for wellbeing and adaptation in semi-arid Africa and India. *World Development*, *125*, 104667. https://doi.org/10.1016/j.worlddev.2019.104667

Reibelt, L., & Nowack, J. (2015). Editorial: Community-based conservation in Madagascar, the 'cure-all' solution? *Madagascar Conservation & Development*, *10*(1), 3. https://doi.org/10.4314/mcd.v10i1.S1

Rogan, M. (2013). Poverty and Headship in Post-apartheid South Africa, 1997–2006. *Social Indicators Research*, *113*(1), 491–511. https://doi.org/10.1007/s11205-012-0107-8

Salerno, J., Borgerhoff Mulder, M., Grote, M. N., Ghiselli, M., & Packer, C. (2016). Household livelihoods and conflict with wildlife in community-based conservation areas across northern Tanzania. *Oryx*, *50*(4), 702–712. https://doi.org/10.1017/S0030605315000393

Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, M., & Reinhardt, J. (2017). Climate change impacts in Sub-Saharan Africa: From physical changes to their social repercussions. *Regional Environmental Change*, *17*(6), 1585–1600. https://doi.org/10.1007/s10113-015-0910-2

Sharaunga, S., Mudhara, M., & Bogale, A. (2015). The Impact of Women's Empowerment in Agriculture on Household Vulnerability to Food Insecurity in the KwaZulu-Natal Province. *Forum for Development Studies*, *42*(2), 195–223. https://doi.org/10.1080/08039410.2014.997792

Sullivan, C., Meigh, J.R., Fediw, T.S. (2002). Derivation and testing of the water poverty index phase 1, Final Report. *Department for International Development, UK*.

Thomas, F. (2007). Global rights, local realities: Negotiating gender equality and sexual rights in the Caprivi Region, Namibia. *Culture, Health & Sexuality*, *9*(6), 599–614. https://doi.org/10.1080/13691050701526263

Tibesigwa, B., & Visser, M. (2016). Assessing Gender Inequality in Food Security among Small-holder Farm Households in urban and rural South Africa. *World Development*, *88*, 33–49. https://doi.org/10.1016/j.worlddev.2016.07.008

Tibesigwa, B., Visser, M., Hunter, L., Collinson, M., & Twine, W. (2015). *Gender Differences in Climate Change Risk, Food Security, and Adaptation: A Study of Rural Households' Reliance on Agriculture and Natural Resources to Sustain Livelihoods*. Environment for Development Centers (EfDC) Discussion Paper (DP) 15-20.

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., Eckley, N., Kasperson, J. X., Luers, A., Martello, M. L., Polsky, C., Pulsipher, A., & Schiller, A. (2003). A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences*, *100*(14), 8074–8079. https://doi.org/10.1073/pnas.1231335100

United Nations Development Programme (UNDP), (2020). *Goal 5: Gender equality | UNDP in Zambia*. Retrieved March 8, 2020, from https://www.zm.undp.org/content/zambia/en/home/sustainable-development-goals/goal-5-gender-equality.html

United States Department of Commerce (USDP), (2020). *ESRL Global Monitoring Division - Global Greenhouse Gas Reference Network.* https://www.esrl.noaa.gov/gmd/ccgg/trends/

World Bank, (2019). *Labor force, female (% of total labor force).* Retrieved from https://data.worldbank.org/indicator/SL.TLF.TOTL.FE.ZS.

World Bank, (1997). *Survey of living conditions: Uttar Pradesh and Bihar. Household Questionnaire, December 1997–March 1998.*

Appendices

Appendix A: Adapted Household Food Insecurity Access Scale (HFIAS) Measurement Questionnaire		
No.	Question	Response Options
1.	Since last June until now, did you worry that your household would not have enough food?	1 = Yes 2 = No
1a.	How often did this Happen?	1 = Rarely (once or twice in the past 12 months) 2 = Sometimes (three to ten times in the past 12 months) 3 = Often (more than ten times in the past 12 months
2.	Since last June until now, were you or any household member not able to eat the kinds of foods you preferred because of a lack of resources?	1 = Yes 2 = No
2a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
3.	Since last June until now, did you or any household member have to eat a limited variety of foods due to a lack of resources?	1 = Yes 2 = No
3a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
4.	Since last June until now, did you or any household member have to eat some foods that you really did not want to eat because of a lack of resources to obtain other types of food?	1 = Yes 2 = No
4a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
5.	Since last June until now, did you or any household member have to eat a smaller meal than you felt you needed because there was not enough food?	1 = Yes 2 = No
5a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
6.	Since last June until now, did you or any household member have to eat fewer meals in a day because there was not enough food	1 = Yes 2 = No

6a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
7.	Since last June until now, was there ever no food to eat any kind in your household because of lack of resources to get food?	1 = Yes 2 = No
7a.	How often did this happen?	1 = Rarely 2 = Sometimes 3 = Often
8.	Since last June until now, did you or any household member go a whole day and night without eating anything because there was not enough food?	1 = Yes 2 = No

Appendix B: Regional P-Values of Grouped Household Head Tests											
Country	Test	Groups	LVI	FC	HWC	SDC	HC	SNC	LDC	WC	ESC
Botswana	T-test	MHH - FHH	0.619	0.0001984 ***FHH	0.35396	0.001577 ***FHH	0.7133	0.01339 ***FHH	3.792e-05 ***MHH	6.461e-07 ***MHH	0.3454
	ANOVA1	All	0.841	5.59e-07 ***	0.887	0.0154 ***	0.327	0.0358 ***	2.42e-08 ***	5.07e-06 ***	0.186
	Tukey HSD	FDJ-FDF	0.9931629	0.0435004 ***FDJ	0.9629338	0.9381442	0.2656755	0.8298736	0.7408495	0.0239970 ***FDF	0.9673716
		MNS-FDF	0.9999708	0.0452213 ***MNS	0.9662861	0.8616556	0.6530846	0.9425600	0.6979380	0.3719684	0.9702869
		MSP-FDF	0.9523129	0.9683627	0.8941636	0.8833294	0.4306749	0.9805329	0.4100938	0.8956427	0.7371739
		MNS-FDJ	0.9788552	0.9105869	0.9996174	0.2351817	0.9567471	0.9969082	0.9824449	0.6955280	0.6044987
		MSP-FDJ	0.9331955	0.0000038 ***FDJ	0.9549531	0.0131468 ***FDJ	0.8916828	0.0225692 ***FDJ	0.0000000 ***MSP	0.0000056 ***MSP	0.5689957
		MSP-MNS	0.8907573	0.0047248 ***MNS	0.9981381	0.9932248	0.9992116	0.5501514	0.0024490 ***MSP	0.4100567	0.2172236
Namibia	T-test	MHH - FHH	0.03364 ***FHH	0.5025	0.02068 ***FHH	0.8886	0.679	0.4583	0.06555	0.3175	0.2603
	ANOVA1	All	0.00208 **	0.00459 ***	0.127	0.0574	0.0341 ***	0.813	0.061	0.211	0.589
	Tukey HSD	FDJ-FDF	0.0389152 ***FDJ	0.9195700	0.9595521	0.0345355 ***FDJ	0.9101247	0.9999647	0.1307375	0.7907385	0.8960871
		MNS-FDF	0.0729501	0.0435081 ***FDF	0.8323078	0.0989903	0.6431944	0.9997209	0.8981212	0.9487370	0.9999999
		MSP-FDF	0.4089135	0.9638350	0.5530233	0.0711132	0.8727857	0.9830094	0.2693164	0.9984769	0.9905809
		MNS-FDJ	0.9885226	0.0048045 ***FDJ	0.9009756	0.9999354	0.0309186 ***MNS	0.9967866	0.2771755	0.9827584	0.7797349

		MSP-FDJ	0.0157857 ***FDJ	0.9752819	0.1203147	0.8987090	0.9933320	0.8211907	0.7759081	0.1750969	0.7074913
		MSP-MNS	0.2808037	0.0027696 ***MSP	0.9765405	0.9787039	0.0236984 ***MNS	0.9215497	0.5606879	0.9191246	0.9770588
Zambia	T-test	MHH - FHH	0.1185	0.1133	0.6778	0.0003724 ***FHH	0.578	0.02052 ***FHH	0.7666	0.4286	0.0007401 ***MHH
	ANOVA1	All	0.423	0.167	0.109	0.00164 ***	0.408	0.0134	0.581	0.1910965	0.00323 ***
	Tukey HSD	FDJ-FDF	0.9996176	0.5043611	0.1686841	0.5700829	0.4622284	0.5639753	0.8810913	0.1910965	0.4630778
		MNS-FDF	0.9977177	0.8230466	0.1228648	0.0975381	0.4164455	0.6835279	0.9979045	0.5564467	0.9248070
		MSP-FDF	0.8329442	0.9962230	0.2063036	0.0099428 ***FDF	0.4159324	0.9736035	0.6632692	0.4695311	0.9058142
		MNS-FDJ	0.9879839	0.9897742	0.9045949	0.9999354	0.9704916	0.9997974	0.9584096	0.9845250	0.9087340
		MSP-FDJ	0.418285	0.1600091	0.9640840	0.0303400 ***FDJ	0.9996473	0.0177326 ***FDJ	0.9510804	0.6387820	0.0018986 ***MSP
		MSP-MNS	0.9394575	0.7863727	0.7308658	0.9999994	0.9475691	0.2343730	0.8192631	0.9873928	0.4479208

**** Indicates significance followed by the abbreviation for the more vulnerable group*

1 All ANOVA results interpreted with consideration to the F-statistic

Appendix C: Complete R-Code

```
---
title: "CBC.LVI"
author: "Andrea Clement"
date: "12/30/2019"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

#Project Title: Assessing Gendered Climate Related Vulnerabilities in Community Based Conservation Areas of Southern Africa

#Research Questions:
> RQ1. What significant differences are present in a Livelihood Vulnerability Index (LVI) of male and female households in the study area?
> RQ2. To what extent might employment by community based conservation areas (CBC's) buffer expected disparities?

#RQ1:
> Introduction: The Livelihood vulnerability index (LVI) is made up of 8 major components. These are listed below with their subcomponets. A formula derived from literature is used to create a overall index that indicates vulnerability faced by each household.

#Major and Sub Components:
> Food: Food insecurity index, crop diversity index
> Wealth: Material wealth index, overall annual income
> Socio-Demographic: Education of household head, proportion of dependants
> Health: Average of reported household health
> Social Networks: Credit use, govermental and conservancy support, participation in community institutions
> Water: Drinking water travel time, drinking water source
> Livelihood Diversity: Livelihood Diversification index, natural resource and agricultural reliance
> Environmental Shock: Agricultural loss due to climate and wildlife

#Methodology for RQ1
```

````
> 0. Load in Libraries and Main Dataframe CSV
> 1. Calculate subcomponents of LVI
> 2. Calculate major components of LVI
> 3. Calculate LVI
> 4. Perform t-tests between LVI scores of male and female headed households
> 6. Perform ANOVA and Tukey HSD tests between LVI scores of FDJ, FDF, MNS, and MSP households
> 8. Perform country-grouped t-tests between LVI scores of male and female heded households
> 9. Perform country-grouped ANOVA and TUKEY HSD tests between LVI scores of FDJ, FDF, MNS, and MSP households
> 6. Data Visualization

#RQ2:
> Introduction: Employment by entities related to the Community Based Conservation Area's (CBCs) may impact many subcomponents used to calculate the LVI. This includes income, material wealth, and support from related organizations. To explore this, employment will be broken down by category and gender. In addition, participation in and support from conservation and governmental entities will be tested for significant differences between male and female headed households.

#Methodology for RQ2:
> 1. Characterize employment in study group
> 2. Other tests related to CBC participation and involvement

-----

#0. Load in libraries and Main Dataframe CSV

```{r Libraries}
#Libraries
library(stringr)
library(plyr)
library(dplyr)
library(car)
library(ggplot2)
library(gplots)
library(fmsb)
library(ggpubr)
```



```{r Main Dataframe}
#Load in main dataframe
````
````

```
data <-read.csv("Full.Data.Master.csv", header=TRUE, na.strings=c("", "97", "98", "99"),
stringsAsFactors = FALSE)
#Fix questionnaire number's 97, 98, 99 (changed to NA with previous code chunk)
data[97,1] <- 97
data[98,1] <- 98
data[99,1] <- 99
data$HH.Gender <- data$Q1GenderofHHhead #rename gender of HH head variable
#Recode gender to 0- male, 1- female

```

```
-----



#RQ1 Part 1. Calculate subcomponents of the major components of LVI

---

#1.1 Food Component (FC)
> A. Food insecurity index
> B. Crop diversity index

```{r Create FC dataframe}

#Create FC dataframe that includes country and gender
FC <- select(data, Questionnaire, Country, HH.Gender)
```

#1.1.A: Food Insecurity Index (FII)
Note: Adapted from collegue who used HFIAS protocol
> 1. Subset data and other prep
> 2. Data cleanup
> 3. Condense two part questions into one variable
> 4. Assign HFIAS protocol calculations
> 5. Subset variable into FC dataframe and write out FII csv

```{r FII.1 Subset data and other prep}

#Run list of vars to pull
vars <- c(
"Country",
"VillageArea",
"JitCoordsDDS",
"JitCoordsDDE",
```
```

```
"Q4PastyearnotenoughfoodYN",
"Q4.1notenoughfoodHowoften",
"Q6PastyearnotabletoeatpreferredfoodYN",
"Q6.1Pastyearnotabletoeatpreferredfoodhowoften",
"Q7PastyearlimitedvarietyfoodYN",
"Q7.1Pastyearlimitedvarietyfoodhowoften",
"Q8PastyeardislikedfoodYN",
"Q8.1Pastyeardislikedfoodhowoften",
"Q9PastyearsmalleramtsfoodYN",
"Q9.1Pastyearsmalleramtsfoodhowoften",
"Q10PastyearFewermealsYN",
"Q10.1PastyearFewermealshowoften",
"Q11PastyearNofoodYN",
"Q11.1PastyearNofoodhowoften",
"Q12PastyearsleephungryYN",
"Q12.1Pastyearsleephungryhowoften",
"Q13PastyearwholedayampnightnofoodYN",
"Q13.1Pastyearwholedayampnightnofoodhowoften")

#Assign variables to dataframe
d <- data[, vars]

#Write a function for future use
proportion <- function(thing) {
        yes <- length(thing[thing==1 & !is.na(thing)])
        n <- length(thing[!is.na(thing)])
        prop <- yes/n
        return(prop)}
```
```

```{r FII.2 Data cleanup}

#Identifiers and indexing
#Name Countries
d$site <- rep("Botswana", nrow(d))
d$site[d$Country==2] <- "Namibia"
d$site[d$Country==3] <- "Zambia"

#Subset FI process variables
f <- c("Q4PastyearnotenoughfoodYN",
"Q4.1notenoughfoodHowoften",
"Q6PastyearnotabletoeatpreferredfoodYN",
"Q6.1Pastyearnotabletoeatpreferredfoodhowoften",
```
```

```
"Q7PastyearlimitedvarietyfoodYN",
"Q7.1Pastyearlimitedvarietyfoodhowoften",
"Q8PastyeardislikedfoodYN",
"Q8.1Pastyeardislikedfoodhowoften",
"Q9PastyearsmalleramtsfoodYN",
"Q9.1Pastyearsmalleramtsfoodhowoften",
"Q10PastyearFewermealsYN",
"Q10.1PastyearFewermealshowoften",
"Q11PastyearNofoodYN",
"Q11.1PastyearNofoodhowoften",
"Q12PastyearsleephungryYN",
"Q12.1Pastyearsleephungryhowoften",
"Q13PastyearwholedayampnightnofoodYN",
"Q13.1Pastyearwholedayampnightnofoodhowoften")

#Assign relevant variables to new dataframe
f <- d[, f]

#Rename question categories
newnames <- c("shortage", "preferred", "variety", "disliked", "amount", "fewmeals", "nofood",
"sleep", "daynight")

```

```{r FII.3 Condense two part questions into one variable}

#Define questions:
#Question pt 1. Yes or no (Example: Past year not enough food?)
#Question pt 2. How often? (Example: Past year not enough food how often?)

#Assign 0 values to households who answered NO to question pt. 1
#Assign 1-3 values to household's who answered YES to question pt. 1, based on "how often"
follow-up question pt. 2

#Scale up the vars and apply no as 1
f$shortage <- f$Q4.1notenoughfoodHowoften #apply original question pt. 2 to new variable
f$shortage[f$Q4PastyearnotenoughfoodYN==2] <- 0 #apply original question pt.1 to same
new variable
table(f$shortage) #vizualize results
f$shortage[f$shortage==22] <- 2 #correct mistake, 22 entered instead of 2

f$preferred <- f$Q6.1Pastyearnotabletoeatpreferredfoodhowoften
f$preferred[f$Q6PastyearnotabletoeatpreferredfoodYN==2] <- 0
table(f$preferred)
```

```
f$variety <- f$Q7.1Pastyearlimitedvarietyfoodhowoften
f$variety[f$Q7PastyearlimitedvarietyfoodYN==2] <- 0
table(f$variety)

f$disliked <- f$Q8.1Pastyeardislikedfoodhowoften
f$disliked[f$Q8PastyeardislikedfoodYN==2] <- 0
table(f$disliked)

f$amount <- f$Q9.1Pastyearsmalleramtsfoodhowoften
f$amount[f$Q9PastyearsmalleramtsfoodYN==2] <- 0
table(f$amount)

f$fewmeals <- f$Q10.1PastyearFewermealshowoften
f$fewmeals[f$Q10PastyearFewermealsYN==2] <- 0
table(f$fewmeals)

f$nofood <- f$Q11.1PastyearNofoodhowoften
f$nofood[f$Q11PastyearNofoodYN==2] <- 0
table(f$nofood)

f$sleep <- f$Q12.1Pastyearsleephungryhowoften
f$sleep[f$Q12PastyearsleephungryYN==2] <- 0
table(f$sleep)

f$daynight <- f$Q13.1Pastyearwholedayampnightnofoodhowoften
f$daynight[f$Q13PastyearwholedayampnightnofoodYN==2] <- 0
table(f$daynight)

#Assign all new created and modified variables to a new dataframe - fi
fi <- f[, newnames] #this will retain 0-3 scale that was applied
head(fi)
```

```{r FII.4 Apply the HFIAS calculations}

#This section uses HFIAS protocol to assign each household a food security status. Categories are: Secure, Slightly Food Insecure (FI), Moderately Food Insecure, and Food Insecure.

#"HFIAS Score (0-27).  Sum of the frequency-of-occurrence during the past four weeks for the 9 food insecurity-related conditions. Sum frequency-of-occurrence question response code (Q1a + Q2a + Q3a + Q4a + Q5a + Q6a + Q7a + Q8a + Q9a)"
#to translate, this is just the sum of the applied 0-3 ordinal
```

```
fi$fiNum <- apply(fi, 1, sum)
hist(fi$fiNum)  #data vizualization, identified 301 value mistake, fixed up top under clean ups
d$Q10.1PastyearFewermealshowoften

fi$fiCat <- rep(NA, nrow(d))
#protocol from from p21 of HFIAS
#if [(shortage=0 or shortage=1) and preferred=0 and variety=0 and disliked=0 and amount=0
and fewmeals=0 and nofood=0 and sleep=0 and daynight=0]
temp <- rep(0, nrow(d))
temp[fi$shortage==0 | fi$shortage==1] <- 1  #quick workaround
fi$fiCat[temp==1 & fi$preferred==0 & fi$variety==0 & fi$disliked==0 & fi$amount==0 &
fi$fewmeals==0 & fi$nofood==0 & fi$sleep==0 & fi$daynight==0] <- "Secure"

#if [(shortage=2 or shortage=3 or preferred=1 or preferred=2 or preferred=3 or variety=1 or
disliked=1) and amount=0 and fewmeals=0 and nofood=0 and sleep=0 and daynight=0]
temp2 <- rep(0, nrow(d))
temp2[fi$shortage==2 | fi$shortage==3 | fi$preferred==1 | fi$preferred==2 | fi$preferred==3
| fi$variety==3 | fi$disliked==1] <- 1  #quick workaround
fi$fiCat[temp2==1 & fi$amount==0 & fi$fewmeals==0 & fi$nofood==0 & fi$sleep==0 &
fi$daynight==0] <- "Slightly FI"

#if [(variety=2 or variety=3 or disliked =2 or disliked =3 or amount=1 or amount=2 or
fewmeals=1 or fewmeals=2) and nofood=0 and sleep=0 and daynight=0]
temp3 <- rep(0, nrow(d))
temp3[fi$variety==2 | fi$variety==3 | fi$disliked==2 | fi$disliked==3 | fi$amount==1 |
fi$amount==2 | fi$fewmeals==1 | fi$fewmeals==2] <- 1  #quick workaround
fi$fiCat[temp3==1 & fi$nofood==0 & fi$sleep==0 & fi$daynight==0] <- "Moderately FI"

#if [amount=3 or fewmeals=3 or nofood=1 or nofood=2 or nofood=3 or sleep=1 or sleep=2 or
sleep=3 or daynight=1 or daynight=2 or daynight=3]
fi$fiCat[fi$amount==3 | fi$fewmeals==3 | fi$nofood==1 | fi$nofood==2 | fi$nofood==3 |
fi$sleep==1 | fi$sleep==2 | fi$sleep==3 | fi$daynight==1 | fi$daynight==2 | fi$daynight==3] <-
"Severely FI"

#Food security levels changed for this project so that lower numbers (1) indicate secure and
higher numbers (4) indicate insecure
d$fiNum <- fi$fiNum
d$fiCat <- ordered(fi$fiCat, levels=c("Secure", "Slightly FI", "Moderately FI", "Severely FI"))
d$fsLev <- rep(NA, nrow(d))
	d$fsLev[d$fiCat=="Secure"] <- 1
	d$fsLev[d$fiCat=="Slightly FI"] <- 2
	d$fsLev[d$fiCat=="Moderately FI"] <- 3
	d$fsLev[d$fiCat=="Severely FI"] <- 4
table(d$fsLev)
```

```
```


```{r FII.6 Subset variable into FC dataframe and write out FII csv}

#Prepare a dataframe containing only questionaire number, site, fiNum (food insecurity number), fiCat (food insecurity category), fsLev (food security level)

#Add new variable "site" to original dataframe
data$Site <- d$site

#Subset dataframe of questionaire number and jittered coordinates
quest <- select(data, Questionnaire, Country, JitCoordsDDS, JitCoordsDDE)
d <- merge(d, quest) #merge with d dataset
food.insecurity <- select(d, Questionnaire, Country, fiNum, fiCat, fsLev) #subset necessary variables
food.insecurity <- food.insecurity[order(food.insecurity$Questionnaire),] #order by questionnaire number

write.csv(food.insecurity, 'food.insecurity.csv')

#Subset fsLev into FC dataframe
FC$FII <- food.insecurity$fsLev
```

#1.1.B: Crop Diversity Index (CDI)

```{r Crop Diversity Index - CDI}

#Calculate Crop Diversity Index
#Inverse of (Number of crop catagories +1)

#Crop catagories (7): maize, sorghum, millet, beans/peanuts, pumpkins/melons, sweet reed, other.

CDI <- select(data, Questionnaire, Q40.1AvgyearMaizeplantedYN,
Q40.1AvgyearSorghumplantedYN, Q40.1AvgyearMilletplantedYN,
Q40.1AvgyearBeanspeanutsplantedYN, Q40.1AvgyearPumpkinsmelonsplantedYN,
Q40.1Avgyearsweet.reedplantedYN, Q40.1AvgyearOTHER2plantedYN)

#Recode so that growing crop category is indicated by a 1, not growing category is indicated by a 0
CDI[2:8] <- ifelse(CDI[2:8] == 1, 1, 0)
```

```
#Total of crop catagories per household
CDI$total <- rowSums(CDI[2:8], na.rm=T)

#Create the index
CDI$CDI <- (1/(CDI$total + 1))

#Assign to FC dataframe
FC$CDI <- CDI$CDI

```

```
---

#1.2 Wealth Component (WC)
> A. Material wealth index
> B. Overall annual income

```{r Create HWC dataframe}

#Create HWC dataframe that includes country and gender

HWC <- select(data, Questionnaire, Country, HH.Gender)

```

```
#1.2.A Material Wealth Index (MWI)
Note: This process was adapted from DHS USAID protocol
> 1. Subset data
> 2. Recode binoimal material assests
> 3. Recode multicatagorical variables
> 5. Recode for land ownership
> 6. Determine number of residents per sleeping rooms
> 7. Subset data to create relative wealth index by country
> 8. Run the PCA's on subsetted country dataframes
> 9. Integrate PCA results into subsetted country dataframes and classify into quintiles
> 10. Rejoin the three subsetted dataframes and visualize results
> 11. Subset variable into HWC dataframe and write out MWI csv


```{r MWI.1 Subset data}

DHSWI <- select(data, Questionnaire, Q3RelationtoHHheadP1, Q3RelationtoHHheadP2,
Q3RelationtoHHheadP3, Q3RelationtoHHheadP4, Q3RelationtoHHheadP5,
```

```
Q3RelationtoHHheadP6, Q3RelationtoHHheadP7, Q3RelationtoHHheadP8,
Q3RelationtoHHheadP9, Q3RelationtoHHheadP10, Q3RelationtoHHheadP11,
Q3RelationtoHHheadP12, Q19Bicycle, Q19Cart, Q19Vehicle, Q19Motorbike, Q19Tractor,
Q19DugoutMokoro, Q19Sledge, Q19Wheelbarrow, Q19Waterpump, Q19Generator,
Q19Solarpanel, Q19TVset, Q19DVDplayer, Q19SatelliteDSTV, Q19Radio, Q19Stove,
Q19FridgeFreezer, Q19Microwaveoven, Q19Electriciron, Q19Coaliron, Q19Cellphone,
Q19Computer, Q19Sewingmaching, Q19Hoe, Q19Plough, Q19Yokes, Q19Rake, Q19Spade,
Q19Axe, Q19Shotgun, Q19Rifle, Q19Spear, Q19Fishingnet, Q19Pots, Q19Jerrycans,
Q19Metaldrums, Q19Mosquitonet, Q21numofsleepingrooms, Q22Roofingmaterial,
Q23Wallmaterial, Q24Floormaterial, Q25Watersource, Q26Toilettype, Q27MainselectricityYN,
Q28YardfencedYN, Q29Fuelusedforcooking, Q30ResidentialpropertywriitendocumentYN,
Q30Residentialpropertytypeofdocument, Q31SizeofResidentialproperty,
Q32Croplandtypeofdocument, Q32CalcTotalcroplandsize, Q41.1Totalcattlekept,
Q41.1Totalgoatskept, Q41.1Totalsheepkept, Q41.1Totaldonkeyskept, Q41.1Totalhorseskept,
Q41.1Totalpoultrykept, Q41.1Totalpigskept, Country)

#Add in country names to dataset
DHSWI$Country <- data$Site
```

```
```{r MWI.2 Recode binomial material assets}

#Material assets originally listed as 1 to indicate the presence of and 2 to indicate the absence of. Recoding all 2's as 0's in indicate the absence of a material item.

#Recode Bicycle (column 4) through mosquito net (column 50)
DHSWI[ , 14:50]  <- lapply(DHSWI[ ,14:50] , FUN = function(x) recode(x, "2=0; 1=1; NA=NA"))

#Recode mains electricity (column 57) and yard fenced (column 58)
DHSWI[ , 57:58]  <- lapply(DHSWI[ ,57:58] , FUN = function(x) recode(x, "2=0"))

```

```
```{r MWI.3 Recode multicatagorical variables}

#Recode multicatagorical variables (fuel source, drinking water, toilet type, floor material, wall material) to binomial variables representing 0/1


#Gather data about multicatagorical variables
#Frequencies of the counts of each data-type observed to make informed decisions about how to group the variablees into binomial categories
count(DHSWI, Q26Toilettype)
```

```
count(DHSWI, Q22Roofingmaterial)
count(DHSWI, Q23Wallmaterial)
count(DHSWI, Q24Floormaterial)
count(DHSWI, Q25Watersource)
count(DHSWI, Q29Fuelusedforcooking)


#Recode toilet type into two indicator variables
#HHToiletOwn: 0 - Use bush or neighbors (3,5), 1 - Have latrine (1,2,4)
#HHToiletTech: 0 - low tech latrine, bush, neighbors (1,3,5); 1 - ventilation improved pit latrine
or flush toilet (2,4)
toilet = DHSWI$Q26Toilettype #variable assignment
DHSWI$HHToiletOwn <- ifelse(toilet == 1| toilet == 2| toilet == 4, 1, 0) #Recode into a new
variable 'ToiletOwn'
DHSWI$HHToiletTech <- ifelse(toilet == 2| toilet == 4, 1, 0)


#Recode Roofing material into indicator variable
#HHRoof: 0- Natural Material or other (1,4); 1- corregated iron, tiles (2,3)
roof = DHSWI$Q22Roofingmaterial
DHSWI$HHRoof <- ifelse(roof == 2| roof == 3, 1, 0)


#Recode wall material into indicator variable
#HHWalls: 0 - Natural materials and other (2:6); 1 - constructed material (1 cement blocks)
wall = DHSWI$Q23Wallmaterial
DHSWI$HHWalls<- ifelse(wall == 1, 1, 0)


#Recode floor material into indicator variable
#HHFloor: 0 - Natural materials and other (2:5); 1 - constructed material (1 cement)
floor = DHSWI$Q24Floormaterial
DHSWI$HHFloors<- ifelse(floor == 1, 1, 0)


#Recode water source into two indicator variables
#HHWaterClosedSource: 0 - Open source (3 river); 1 - Closed source (1,2,4,5,6)
#HHWaterPrivate: 0 - Shared source (2-6); 1 - Private mains connection (1)
water = DHSWI$Q25Watersource
DHSWI$HHWaterClosedSource <- ifelse(water ==1| water==2| water==4| water==5|
water==6, 1, 0)
DHSWI$HHWaterPrivate <- ifelse(water == 1, 1, 0)
```

```
#Recode fuel type into indicator variable
#HHFuel: 0 - gathered (firewood 1), 1 - purchased (2:4)
fuel = DHSWI$Q29Fuelusedforcooking
DHSWI$HHFuels<- ifelse(fuel == 1, 0, 1)


#Summary of data to reference to original dataset and ensure accuracy
summary(DHSWI$HHFuels)
summary(DHSWI$HHWaterClosedSource)
```

```{r MWI.4 Recode for house ownership}

#Observe frequencies of the counts of each data-type observed to make informed decisions about how to group the variables into binomial categories
count(DHSWI, Q30Residentialpropertytypeofdocument)

#Recode 2-land board certificate, 4-title deed, 5-khuta/kgatla letter as 1 #Recode 98-missing data, 1-rental lease agreement, 3-land board lease agreement as 0
House = DHSWI$Q30Residentialpropertytypeofdocument #variable assignment
DHSWI$HouseOwn <- ifelse(House == 2| House == 4| House == 5, 1, 0) #recode into a new variable 'HouseOwn'

#Summary of data to reference to original dataset and ensure accuracy
summary(DHSWI$HouseOwn)
count(DHSWI, HouseOwn)
```


```{r MWI.5 Recode for land ownership}

#Frequencies of the counts of each data-type observed to make informed decisions about how to group the variablees into binomial catagories
count(DHSWI, Q32Croplandtypeofdocument)

#Recode 2-land board certificate, 4-title deed, 5-khuta/kgatla letter as 1 #Recode 98-missing data, 1-rental lease agreement, 3-land board lease agreement as 0
Land = DHSWI$Q32Croplandtypeofdocument
DHSWI$LandOwn <- ifelse(Land == 2| Land == 4| Land == 5, 1, 0)

#Summary of data to reference to original dataset and ensure accuracy
summary(DHSWI$LandOwn)
count(DHSWI, LandOwn)
```

```
```{r MWI.6 Determine number of residents per sleeping room}

#Determine number of residents

#Recode resident fields to read as 1 for the prescence of a person (previously coded values 1-9
determing household relation) or 0 for the absence of a person (previously coded NA for not
applicable)

#Recode mistake value of 33 as 3 in column 12
DHSWI$Q3RelationtoHHheadP11 <- recode(DHSWI$Q3RelationtoHHheadP11, "33=3")

#Change all values from 1-9 in columns 2 through 13 to 1.
DHSWI[ , 2:13]  <- lapply(DHSWI[ ,2:13] , FUN = function(x) recode(x, "1:9=1"))

#Summary of data to reference to original dataset and ensure accuracy
sum(DHSWI$Q3RelationtoHHheadP2, na.rm = T)
sum(DHSWI$Q3RelationtoHHheadP11, na.rm = T)
sum(DHSWI$Q3RelationtoHHheadP12, na.rm = T)

#Create a new column with the sum of residents using RelationtoHHhead columns 1-12

#Add together how many residents present in each household and assign to a new variable
DHSWI$sumResidents <- rowSums(DHSWI[2:13], na.rm=T)

#Summary of data to reference to original dataset and ensure accuracy
summary(DHSWI$sumResidents)
head(DHSWI$sumResidents)

#Determine number of Sleeping Rooms by sumResidents

#Assign NA value to households in which sleeping rooms = 0 to avoid infinate values
DHSWI$Q21numofsleepingrooms[DHSWI$Q21numofsleepingrooms == 0] <- NA

#Divide number of household rooms by the number of residents and assign to a
DHSWI$resPerRoom <- DHSWI$sumResidents/DHSWI$Q21numofsleepingrooms

summary(DHSWI$resPerRoom)
```


```{r MWI.7 Subset Variables for PCA}

#Subset Relevant Variables for Principle Component Analysis (PCA)
```
```

```
#Name variables to be subsetted
#Pulled out horses variable because of no variation
vars <- c("resPerRoom", "LandOwn", "HouseOwn", "Q19Cart", "Q19Vehicle", "Q19Motorbike",
"Q19Tractor", "Q19DugoutMokoro", "Q19Sledge", "Q19Wheelbarrow", "Q19Waterpump",
"Q19Generator", "Q19Solarpanel", "Q19TVset", "Q19DVDplayer", "Q19SatelliteDSTV",
"Q19Radio", "Q19Stove", "Q19FridgeFreezer", "Q19Microwaveoven", "Q19Electriciron",
"Q19Coaliron", "Q19Cellphone", "Q19Computer", "Q19Sewingmaching", "Q19Hoe",
"Q19Plough", "Q19Yokes", "Q19Rake", "Q19Spade", "Q19Axe", "Q19Shotgun", "Q19Rifle",
"Q19Spear", "Q19Fishingnet", "Q19Pots", "Q19Jerrycans", "Q19Metaldrums",
"Q19Mosquitonet", "Q27MainselectricityYN", "Q28YardfencedYN",
"Q31SizeofResidentialproperty", "Q32CalcTotalcroplandsize", "Q41.1Totalcattlekept",
"Q41.1Totalgoatskept", "Q41.1Totaldonkeyskept", "Q41.1Totalsheepkept",
"Q41.1Totalpigskept", "Q41.1Totalpoultrykept", "HHToiletOwn", "HHToiletTech", "HHRoof",
"HHWalls", "HHFloors", "HHWaterClosedSource", "HHWaterPrivate", "HHFuels",
"sumResidents", "Q19Bicycle")

#Subset data into a new dataframe with NA values recoded as 0's titled 'subset.wi.noNA'
subset.wi.noNA <- DHSWI[vars]
subset.wi.noNA[is.na(subset.wi.noNA)] <- 0 #recode NA values as 0


#Subset same variables with Questionnaire in order to merge at the end
vars2 <- c("Questionnaire", "resPerRoom", "LandOwn", "HouseOwn", "Q19Cart",
"Q19Vehicle", "Q19Motorbike", "Q19Tractor", "Q19DugoutMokoro", "Q19Sledge",
"Q19Wheelbarrow", "Q19Waterpump", "Q19Generator", "Q19Solarpanel", "Q19TVset",
"Q19DVDplayer", "Q19SatelliteDSTV", "Q19Radio", "Q19Stove", "Q19FridgeFreezer",
"Q19Microwaveoven", "Q19Electriciron", "Q19Coaliron", "Q19Cellphone", "Q19Computer",
"Q19Sewingmaching", "Q19Hoe", "Q19Plough", "Q19Yokes", "Q19Rake", "Q19Spade",
"Q19Axe", "Q19Shotgun", "Q19Rifle", "Q19Spear", "Q19Fishingnet", "Q19Pots",
"Q19Jerrycans", "Q19Metaldrums", "Q19Mosquitonet", "Q27MainselectricityYN",
"Q28YardfencedYN", "Q31SizeofResidentialproperty", "Q32CalcTotalcroplandsize",
"Q41.1Totalcattlekept", "Q41.1Totalgoatskept", "Q41.1Totaldonkeyskept",
"Q41.1Totalsheepkept", "Q41.1Totalpigskept", "Q41.1Totalpoultrykept", "HHToiletOwn",
"HHToiletTech", "HHRoof", "HHWalls", "HHFloors", "HHWaterClosedSource",
"HHWaterPrivate", "HHFuels", "sumResidents", "Q19Bicycle")

#Assign second variable set to dataframe with questionaire variable
subset.wi.quest <- DHSWI[vars2]
subset.wi.quest[is.na(subset.wi.quest)] <- 0 #recode NA values as 0
```

```
```{r MWI.8 Run PCA}

#PCA is ran with derived eigenvectors (weights) from the correlation matrix
#Correlation matrix is used instead of covariance matrix because the information is not
standardized
#Without this step the quantitative variables would dominate binary variables
pca.cor <-princomp(na.omit(subset.wi.noNA), cor=TRUE)

#Evaluate the PCA results using various functions
summary(pca.cor)
loadings(pca.cor)

#Visualize the PCA
plot(pca.cor, main = "PCA Components by Percent Variance")
biplot(pca.cor, xlab = "Component 1", ylab = "Component 2", main = "


Interaction of PCA Components 1 and 2



")

```

```{r MWI.9 Integrate PCA results into dataframe and classify into quintiles}

#J. Integrate PCA results into dataframe

#Visualize the scores of each principle component
pca.cor$scores

#Create new dataframe titled 'scores' out of principle component scores
scores <- data.frame(pca.cor$scores)

#Visualize the head of the new dataframe 'scores'
head(scores)

#Create a new variable titled 'PC1' out of the scores of the first principle component
subset.wi.noNA$PC1 <- scores$Comp.1

#Classify PC1 Results into quintiles
```
```

```
#Scores of PC1 are used to represent relative wealth
#PC1 scores are split into quintiles to assign each household a weatlh index rating
#Scores listed under 'quint' variable represent 1 as the poorest and 5 as the wealthiest quintile

#Create a new variable called 'quintile' from the 'subset.wi.noNA' dataframe containing
quintile status of PC1 scores. Recode into a new dataframe named 'quint'
subset.wi.noNA %>% mutate(quintile = ntile(PC1, 5)) -> quint

```

```{r MWI.10 Visualize results}

#Visualize the new dataframe to ensure accuracy
head(quint)

#Merge with subset.wi.quest dataframe, which contains questionnaire number. Questionnaire
number was pulled out in order to prevent skewing of PCA results
quint <- merge(subset.wi.quest, quint)

#Visualize the data with simple plot function'
plot(x = quint$Questionnaire, y = quint$PC1, col = quint$quintile, main = 'Wealth Index Score
by Household', ylab = 'Index Score',
    xlab = 'Household ID
    (0-240 = Namibia, 241-481 = Botswana, 482-726 = Zambia)')+
abline(v = 240, col = 'gray') +
abline(v = 481, col = 'gray') #adding markers at lines 240 and 481

#Visualize the data with ggplot
ggplot(data= quint, mapping = aes(x= Questionnaire, y =PC1, color = quintile)) + geom_point()
+ labs(x= 'Household ID
(Namibia: 0-240, Botswana: 241-481, Zambia: 482-726)', y = 'Index Score', title = "Wealth
Index Score by Household") + geom_vline(xintercept = 240, na.rm = FALSE, show.legend =
NA)+ geom_vline(xintercept = 481, #adding line at markers 240 and 481
  na.rm = FALSE, show.legend = NA) +theme(plot.title = element_text(size =18, hjust = 0.8))
#centering and enlarging title

```

```{r MWI.11 Subset variable into HWC dataframe and write out MWI csv}

#Subset out calculated variables
MWI <- select(quint, Questionnaire, resPerRoom, sumResidents, PC1, quintile)

#Write out CSV
```

```
write.csv(MWI, 'MWI')

HWC$MWI <- MWI$quintile

```
```

#1.2.B Overall Annual Income (INC)
Note: OAI includes monatary value of all agricultural and resource products, as well as total cash income
> 1. Subset variables
> 2. Calculate average prices for agriculture
> 3. Calculate average prices for natural resources
> 4. Replace missing country average prices
> 5. Calculate total cash income per household in US dollars
> 6. Calculate total cash values of all agriculture and natural resouce materials obtained by household
> 7. Create totals dataframe
> 8. Calculate total cash values
> 9. Calculate income percentages
> 10. Subset variable into HWC dataframe and write out OAI csv

```{r INC.1 Subset Variables}
#Subset Variables for household totals of income, agriculture, and natural resouces

#Select variables related to cash income
cash <- select(data, Questionnaire, ExchangeRatetoUSused, Q14CalcOverallTotal,
Q15Totalamountreceivedinremittancesinpastyear,
Q17CalcTotalamtHHearnedfromlabourprogram,
Q17CalcTotalamtHHreceivedfromgovtpension,
Q17CalcTotalamtHHreceivedfromgovtOVCsupport,
Q17CalcTotalamtHHreceivedfromgovtreliefprogram,
Q17CalcTotalamtHHreceivedfromconservancytrustcashpmt,
Q17CalcTotalamtHHreceivedfromothercashpmt)

#Select variables related to agriculture to include all harvested, livestock info left as cash earned from sales
ag2 <- select(data, Questionnaire, Q39Amtmaizeharvested, Q39AmtSorghumharvested,
Q39AmtMilletharvested, Q39AmtBeansPeanutsharvested,
Q39AmtPumpkinsMelonsharvested, Q39Amtsweet.reedharvested, Q39AmtOTHER2harvested,
Q44Totalcashearnedfromcattlesales, Q44Totalcashearnedfromgoatssales,
Q44Totalcashearnedfromsheepsales, Q44Totalcashearnedfromdonkeyssales,
Q44Totalcashearnedfromhorsessales, Q44Totalcashearnedfrompoultrysales,
Q44Totalcashearnedfrompigsales)
```
```

```
#Select variables related to natural resouces gathered to include all gathered in wet and dry
season
nr.1 <- select(data, Questionnaire, Q50Firewoodwetseasontotalgathered,
Q50Firewooddryseasontotalgathered, Q50Thatchinggrasswetseasontotalgathered,
Q50Thatchinggrassdryseasontotalgathered, Q50Fishwetseasontotalgathered,
Q50Fishdryseasontotalgathered, Q50Papyruswetseasontotalgathered,
Q50Papyrusdryseasontotalgathered, Q50Grapplewetseasontotalgathered,
Q50Grappledryseasontotalgathered, Q50Reedswetseasontotalgathered,
Q50Reedsdryseasontotalgathered, Q50Buildingpoleswetseasontotalgathered,
Q50Buildingpolesdryseasontotalgathered, Q50Mudforbrickswetseasontotalgathered,
Q50Mudforbricksdryseasontotalgathered, Q50Palmleaveswetseasontotalgathered,
Q50Palmleavesdryseasontotalgathered, Q50waterlilywetseasontotalgathered,
Q50waterlilydryseasontotalgathered, Q50Medicinalplantswetseasontotalgathered,
Q50Medicinalplantsdryseasontotalgathered, Q50Edibleplantswetseasontotalgathered,
Q50Edibleplantsdryseasontotalgathered, Q50Birdsandanimalswetseasontotalgathered,
Q50Birdsandanimalsdryseasontotalgathered, Q50Otherresourcewetseasontotalgathered,
Q50Otherresourcedryseasontotalgathered)

#Combine nr wet and dry season totals for natural resources
nr.2 <- select(nr.1, Questionnaire)
nr.2$wood <- nr.1$Q50Firewoodwetseasontotalgathered +
nr.1$Q50Firewooddryseasontotalgathered
nr.2$grass <- nr.1$Q50Thatchinggrasswetseasontotalgathered +
nr.1$Q50Thatchinggrassdryseasontotalgathered
nr.2$fish <- nr.1$Q50Fishwetseasontotalgathered + nr.1$Q50Fishdryseasontotalgathered
nr.2$papyrus <- nr.1$Q50Papyruswetseasontotalgathered +
nr.1$Q50Papyrusdryseasontotalgathered
nr.2$grapple <- nr.1$Q50Grapplewetseasontotalgathered +
nr.1$Q50Grappledryseasontotalgathered
nr.2$reeds <- nr.1$Q50Reedswetseasontotalgathered +
nr.1$Q50Reedsdryseasontotalgathered
nr.2$pole <- nr.1$Q50Buildingpoleswetseasontotalgathered +
nr.1$Q50Buildingpolesdryseasontotalgathered
nr.2$mud <- nr.1$Q50Mudforbrickswetseasontotalgathered +
nr.1$Q50Mudforbricksdryseasontotalgathered
nr.2$palm <- nr.1$Q50Palmleaveswetseasontotalgathered +
nr.1$Q50Palmleavesdryseasontotalgathered
nr.2$lily <- nr.1$Q50waterlilywetseasontotalgathered +
nr.1$Q50waterlilydryseasontotalgathered
nr.2$medplant <- nr.1$Q50Medicinalplantswetseasontotalgathered +
nr.1$Q50Medicinalplantsdryseasontotalgathered
nr.2$edplant <- nr.1$Q50Edibleplantswetseasontotalgathered +
nr.1$Q50Edibleplantsdryseasontotalgathered
```

```
nr.2$bird.anim <- nr.1$Q50Birdsandanimalswetseasontotalgathered +
nr.1$Q50Birdsandanimalsdryseasontotalgathered
nr.2$other <- nr.1$Q50Otherresourcewetseasontotalgathered +
nr.1$Q50Otherresourcedryseasontotalgathered
nr.2$country <- data$Country

```



```{r INC.2 Calculate average prices for agriculture}

#Calculate average prices for each Agriculture material by country in US dollars

#subset data to include total earned and total sold in order to calculate averages
ag.avg <- select(data, Questionnaire, Country, Q39Totalearnedfrommaize,
Q39TotalearnedfromSorghum, Q39TotalearnedfromMillet,
Q39TotalearnedfromBeansPeanuts, Q39TotalearnedfromPumpkinsMelons,
Q39Totalearnedfromsweet.reed, Q39TotalearnedfromOTHER2,
Q44Totalcashearnedfromcattlesales, Q44Totalcashearnedfromgoatssales,
Q44Totalcashearnedfromsheepsales, Q44Totalcashearnedfromdonkeyssales,
Q44Totalcashearnedfromhorsessales, Q44Totalcashearnedfrompoultrysales,
Q44Totalcashearnedfrompigsales, Q39Amtmaizesold, Q39AmtSorghumsold,
Q39AmtMilletsold, Q39AmtBeansPeanutssold, Q39AmtPumpkinsMelonssold,
Q39Amtsweet.reedsold, Q39AmtOTHER2sold, Q44numofcattlesold, Q44numofgoatssold,
Q44numofsheepsold, Q44numofdonkeyssold, Q44numofhorsessold, Q44numofpoultrysold,
Q44numofpigssold)

#aggregate for totals by country of amount sold/amount earned
ag.avg2 <- aggregate(ag.avg[, 3:30], list(ag.avg$Country), mean, na.rm = T)
#create new table to store price per unit values
ag.avg3 <- select(ag.avg2, Group.1)

#find price per unit by dividing average earned by amount sold
ag.avg3$maize.unit <- ag.avg2$Q39Totalearnedfrommaize/ ag.avg2$Q39Amtmaizesold

#repeat for all other products
ag.avg3$sorghum.unit <-
ag.avg2$Q39TotalearnedfromSorghum/ag.avg2$Q39AmtSorghumsold
ag.avg3$millet.unit <- ag.avg2$Q39TotalearnedfromMillet/ag.avg2$Q39AmtMilletsold
ag.avg3$beans.unit <-
ag.avg2$Q39TotalearnedfromBeansPeanuts/ag.avg2$Q39AmtBeansPeanutssold
ag.avg3$pump.unit <-
ag.avg2$Q39TotalearnedfromPumpkinsMelons/ag.avg2$Q39AmtPumpkinsMelonssold
```

```
ag.avg3$swreed.unit <-
ag.avg2$Q39Totalearnedfromsweet.reed/ag.avg2$Q39Amtsweet.reedsold
ag.avg3$other2.unit <- ag.avg2$Q39TotalearnedfromOTHER2/ag.avg2$Q39AmtOTHER2sold
ag.avg3$cattle <- ag.avg2$Q44Totalcashearnedfromcattlesales/ag.avg2$Q44numofcattlesold
ag.avg3$goats <- ag.avg2$Q44Totalcashearnedfromgoatssales/ag.avg2$Q44numofgoatssold
ag.avg3$sheep <- ag.avg2$Q44Totalcashearnedfromsheepsales/ag.avg2$Q44numofsheepsold
ag.avg3$donkeys <-
ag.avg2$Q44Totalcashearnedfromdonkeyssales/ag.avg2$Q44numofdonkeyssold
ag.avg3$horses <-
ag.avg2$Q44Totalcashearnedfromhorsessales/ag.avg2$Q44numofhorsessold
ag.avg3$poultry <-
ag.avg2$Q44Totalcashearnedfrompoultrysales/ag.avg2$Q44numofpoultrysold
ag.avg3$pigs <- ag.avg2$Q44Totalcashearnedfrompigsales/ag.avg2$Q44numofpigssold

#subset exchange rates
exchange <- select(data, Country, ExchangeRatetoUSused)
#aggregate
exchange <- aggregate(exchange, list(exchange$Country), mean)
rate <- exchange$ExchangeRatetoUSused

#convert all values in ag.avg3 to US dollars
ag.US <- select(ag.avg3, Group.1)
ag.US$maize <- ag.avg3$maize.unit/rate
ag.US$sorghum <- ag.avg3$sorghum.unit/rate
ag.US$millet <- ag.avg3$millet.unit/rate
ag.US$beans <- ag.avg3$beans.unit/rate
ag.US$pump <- ag.avg3$pump.unit/rate
ag.US$swreed <- ag.avg3$swreed.unit/rate
ag.US$other2 <- ag.avg3$other2/rate
ag.US$cattle <- ag.avg3$cattle/rate
ag.US$goats <- ag.avg3$goats/rate
ag.US$sheep <- ag.avg3$sheep/rate
ag.US$donkeys <- ag.avg3$donkeys/rate
ag.US$horses <- ag.avg3$horses/rate
ag.US$poultry <- ag.avg3$poultry/rate
ag.US$pigs <- ag.avg3$pigs/rate

#round ag.US to two digits
ag.US <- round(ag.US, 2)
```

```{r INC.3 Calculate average prices for natural resources}

#Calculate average prices for each natural resource material by country in US dollars
```

```
#subset data to include total earned and total sold in order to calculate averages
nr.avg <- select(data, Questionnaire, Q51Quantityfirewoodsoldnpastyear,
Q51Totalearnedfromfirewoodinpastyear, Q51QuantityThatchinggrasssoldnpastyear,
Q51TotalearnedfromThatchinggrassinpastyear, Q51QuantityFishsoldnpastyear,
Q51TotalearnedfromFishinpastyear, Q51Quantitypapyrussoldnpastyear,
Q51Totalearnedfrompapyrusinpastyear, Q51Quantitygrapplesoldinpastyear,
Q51Totalearnedfromgrappleinpastyear, Q51Quantityreedssoldnpastyear,
Q51Totalearnedfromreedsinpastyear, Q51Quantitybuildingpolessoldnpastyear,
Q51Totalearnedfrombuildingpolesinpastyear, Q51Quantitymudforbrickssoldnpastyear,
Q51Totalearnedfrommudforbricksinpastyear, Q51Quantitypalmleavessoldnpastyear,
Q51Totalearnedfrompalmleavesinpastyear, Q51Quantitywaterlilysoldnpastyear,
Q51Totalearnedfromwaterlilyinpastyear, Q51Quantitymedicinalplantssoldnpastyear,
Q51Totalearnedfrommedicinalplantsinpastyear, Q51Quantityedibleplantssoldnpastyear,
Q51Totalearnedfromedibleplantsinpastyear, Q51Quantitybirdsandanimalssoldnpastyear,
Q51Totalearnedfrombirdsandanimalsinpastyear, Q51Quantityotherresourcesoldnpastyear,
Q51Totalearnedfromotherresourceinpastyear)

#aggregate for totals by country of amount sold/amount earned
nr.avg2 <- aggregate(nr.avg[, 2:29], list(ag.avg$Country), mean, na.rm = T)
#create new table to store price per unit values
nr.avg3 <- select(nr.avg2, Group.1)

#find average price per unit of each material
nr.avg3$firewood <-
nr.avg2$Q51Totalearnedfromfirewoodinpastyear/nr.avg2$Q51Quantityfirewoodsoldnpastyea
r
nr.avg3$grass <-
nr.avg2$Q51TotalearnedfromThatchinggrassinpastyear/nr.avg2$Q51QuantityThatchinggrasss
oldnpastyear
nr.avg3$fish <-
nr.avg2$Q51TotalearnedfromFishinpastyear/nr.avg2$Q51QuantityFishsoldnpastyear
nr.avg3$papyrus <-
nr.avg2$Q51Totalearnedfrompapyrusinpastyear/nr.avg2$Q51Quantitypapyrussoldnpastyear
nr.avg3$grapple <-
nr.avg2$Q51Totalearnedfromgrappleinpastyear/nr.avg2$Q51Quantitygrapplesoldinpastyear
nr.avg3$reeds <-
nr.avg2$Q51Totalearnedfromreedsinpastyear/nr.avg2$Q51Quantityreedssoldnpastyear
nr.avg3$poles <-
nr.avg2$Q51Totalearnedfrombuildingpolesinpastyear/nr.avg2$Q51Quantitybuildingpolessold
npastyear
nr.avg3$mud <-
nr.avg2$Q51Totalearnedfrommudforbricksinpastyear/nr.avg2$Q51Quantitymudforbrickssold
npastyear
```

````
nr.avg3$palm <-
nr.avg2$Q51Totalearnedfrompalmleavesinpastyear/nr.avg2$Q51Quantitypalmleavessoldnpastyear
nr.avg3$lily <-
nr.avg2$Q51Totalearnedfromwaterlilyinpastyear/nr.avg2$Q51Quantitywaterlilysoldnpastyear
nr.avg3$mplants <-
nr.avg2$Q51Totalearnedfrommedicinalplantsinpastyear/nr.avg2$Q51Quantitymedicinalplantssoldnpastyear
nr.avg3$edplants <-
nr.avg2$Q51Totalearnedfromedibleplantsinpastyear/nr.avg2$Q51Quantityedibleplantssoldnpastyear
nr.avg3$other <-
nr.avg2$Q51Totalearnedfromotherresourceinpastyear/nr.avg2$Q51Quantityotherresourcesoldnpastyear

#convert all values to us dollars
nr.US <- select(nr.avg3, Group.1)
nr.US <- nr.avg3[,2:14] / exchange[,3]
nr.US$Country <- nr.avg3$Group.1
nr.US <- nr.US[,c(14,1:13)]

#round ag.US to two digits
nr.US <- round(nr.US, 2)

```



```{r INC.4 Replace missing country average prices}

#Replace missing values in nr.US and ag.US dataframes

#Use overall average for countries missing values
#If all countries are missing values, omit from the study or fill in using another source?

#Missing values:
#Country 1: grapple, other, millet, pump, other2, donkeys, horses, pigs
#Country 2: papyrus, mud, lily, mplants, edplants, other, millet, pump, swreed, other2, goats, sheep, donkeys, horses, pigs
#Country 3: papyrus, mud, lily, swreed, other2, sheep, donkeys, horses

#assign a dummy value "1" in order to aggregate nr table
nr.US$group <- 1

#aggregate nr table to find overall natural resouce material totals
````

```
all.nr.US<- aggregate(nr.US[, 2:14], list(nr.US$group), mean, na.rm = T)

#repeat with ag table
#assign a dummy value "1" in order to aggregate ag table
ag.US$group <- 1

#aggregate ag table to find overall agriculture material totals
all.ag.US<- aggregate(ag.US[, 2:15], list(ag.US$group), mean, na.rm = T)

#combine nr/ag tables
ag.US$Country <- ag.US$Group.1
agnr.US <- full_join(ag.US, nr.US, by = "Country")

#replace values
agnr.US[1,4] <- all.ag.US$millet
agnr.US[2,4] <- all.ag.US$millet
agnr.US[1,6] <- all.ag.US$pump
agnr.US[2,6] <- all.ag.US$pump
agnr.US[2,7] <- all.ag.US$swreed
agnr.US[3,7] <- all.ag.US$swreed
agnr.US[2,10] <- all.ag.US$goats
agnr.US[2,11] <- all.ag.US$sheep
agnr.US[3,11] <- all.ag.US$sheep
agnr.US[1,15] <- all.ag.US$pigs
agnr.US[2,15] <- all.ag.US$pigs
agnr.US[2,21] <- all.nr.US$papyrus
agnr.US[3,21] <- all.nr.US$papyrus
agnr.US[1,22] <- all.nr.US$grapple
agnr.US[2,25] <- all.nr.US$mud
agnr.US[3,25] <- all.nr.US$mud
agnr.US[2,27] <- all.nr.US$lily
agnr.US[3,27] <- all.nr.US$lily
agnr.US[2,28] <- all.nr.US$mplants
agnr.US[1,30] <- all.nr.US$other
agnr.US[2,30] <- all.nr.US$other

#After reviewing item values for all countries, it was decided that values for Namibian cattle,
sorghum, palm, and Botswana firewood were skewed to be unusually high due to small
sample sizes. Below, these values are replaced with the average of the other two country
values.

#Namibian cattle
agnr.US[2,9] <- (agnr.US[1,9]+agnr.US[3,9])/2
#Namibian sorghum
```

````
agnr.US[2,3] <- (agnr.US[1,3]+agnr.US[3,3])/2
#Namibian palm
agnr.US[2,9] <- (agnr.US[1,9]+agnr.US[3,9])/2
#Botswana firewood
agnr.US[1,18] <- (agnr.US[2,18]+agnr.US[3,18])/2
```



```{r INC.5 Calculate total cash income per household in US dollars}
#Calculate total cash income per household in US dollars

#add together all cash income and assign to a new variable
#NOTE: Gvt. and conservancy support subtracted later from total
cash$totalcash <- rowSums(cash[3:10], na.rm = T)
#convert to US dollars using the given exchange rate
cash$conv.cash <- cash$totalcash/cash$ExchangeRatetoUSused

```



```{r INC.6 Create totals dataframe}
#Create dataframes of household US cash values of all natural resources and agricultural materials harvested

#merge dataframe containing each households nr/ag and nr/ag totals by country
country <- select(data, Questionnaire, Country)
house.totals <- merge(nr.2, ag2)
house.totals <- merge(house.totals, country)
house.price <- merge(country, agnr.US, by = "Country")

#create df of cash values of nr collected
nrcash <- select(data, Questionnaire)
nrcash$wood <- house.totals$wood * house.price$firewood
nrcash$grass <- house.totals$grass * house.price$grass
nrcash$fish <- house.totals$fish * house.price$fish
nrcash$papyrus <- house.totals$papyrus * house.price$papyrus
nrcash$grapple <- house.totals$grapple * house.price$grapple
nrcash$reeds <- house.totals$reeds * house.price$reeds
nrcash$pole <- house.totals$pole * house.price$poles
nrcash$mud <- house.totals$mud * house.price$mud
nrcash$palm <- house.totals$palm * house.price$palm
nrcash$lily <- house.totals$lily * house.price$lily
nrcash$medplant <- house.totals$medplant * house.price$mplants
nrcash$edplant <- house.totals$edplant * house.price$edplants
````

```
nrcash$bird.anim <- house.totals$bird.anim * 0
nrcash$other2 <- house.totals$other * house.price$other2

#create df of cash values of ag collected
agcash <- select(ag2, Questionnaire)
agcash$maize <- house.totals$Q39Amtmaizeharvested * house.price$maize
agcash$sorghum <- house.totals$Q39AmtSorghumharvested * house.price$sorghum
agcash$millet <- house.totals$Q39AmtMilletharvested * house.price$millet
agcash$beans.pean <- house.totals$Q39AmtBeansPeanutsharvested * house.price$beans
agcash$pump.mel <- house.totals$Q39AmtPumpkinsMelonsharvested * house.price$pump
agcash$sw.reed <- house.totals$Q39Amtsweet.reedharvested * house.price$swreed
agcash$other <- house.totals$Q39AmtOTHER2harvested * house.price$other

#add in animal sales converted to US dollars
ex.US <- cash$ExchangeRatetoUSused
agcash$cattle <- ag2$Q44Totalcashearnedfromcattlesales/ex.US
agcash$goat <- ag2$Q44Totalcashearnedfromgoatssales/ex.US
agcash$sheep <- ag2$Q44Totalcashearnedfromsheepsales/ex.US
agcash$donkey <- ag2$Q44Totalcashearnedfromdonkeyssales/ex.US
agcash$horse <- ag2$Q44Totalcashearnedfromhorsessales/ex.US
agcash$poultry <- ag2$Q44Totalcashearnedfrompoultrysales/ex.US
agcash$pig <- ag2$Q44Totalcashearnedfrompigsales/ex.US
` ` `


` ` `{r INC.7 Calculate total cash values}
#Calculate total cash values of all agriculture and natural resouce materials obtained by household

#Add together totals of agriculture/nr cash values
#add totals together across rows excluding first column containing questionaire number
agcash$totalagcash <- rowSums(agcash[2:15], na.rm = T)
nrcash$totalnrcash <- rowSums(nrcash[2:15], na.rm = T)

` ` `


` ` `{r INC.8 Calculate income percentages}
#Calculate percentage of income from natural resources, and percentage of income from natural resouces and agriculture

#create new totals dataframe with cash, agriculture, and natural resouce totals
totals <- select(cash, Questionnaire, conv.cash)
totals$totalagcash <- agcash$totalagcash
```

````
totals$totalnrcash <- nrcash$totalnrcash

#create variable composed of the total of natural resouces and agriculture
totals$total.ag.nr <- rowSums(totals[3:4])

#create a variable composed of all household income with natural resource and agricultural values added in
totals$total.all <- rowSums(totals[2:4])

#find percentage of income derived from natural resouces
totals$nr.rel <- totals$totalnrcash/totals$total.all
#find percentage of income derived from agriculture
totals$ag.rel <- totals$totalagcash/totals$total.all
#find percentage of income derived from natural resouces and agriculture
totals$nr.ag.rel <- totals$total.ag.nr/totals$total.all


plot(totals$nr.rel~totals$Questionnaire)

country <- select(data, Questionnaire, Country)
totals <- merge(totals, country)

plotmeans(totals$nr.rel ~ totals$Country)
```


```{r INC.9 Subset variable into HWC dataframe and write out INC csv}

HWC$INC <- totals$total.all
write.csv(totals, "INC.csv")
```


---

#1.3 Socio-Demographic Component (SDC):
> A. Education of household head
> B. Percentage of dependants

```{r SDC dataframe}
#Create new dataframe for SDC variables that includes gender and country
SDC <- select(data, Questionnaire, Country, HH.Gender)
```


#1.3.A Education of Household Head (EDU)
````
````

````
```{r EDU - Education of Household Head}
#Subset the ordinal catagories of education for each household head

SDC$EDU <- data$Q3HighesteducationP1

```


#1.3.B Percentage of dependants (DEP)

```{r DEP - Percentage of dependants}
#Calculate number of people in each houshold below 17 and over 65

#Subset data and calculate
dep <- select(data, Questionnaire, Q3AgeP1, Q3AgeP2, Q3AgeP3,  Q3AgeP4, Q3AgeP5,
Q3AgeP6, Q3AgeP7, Q3AgeP8, Q3AgeP9,  Q3AgeP10, Q3AgeP11, Q3AgeP12)

#Recode Q3AgeP1 through QAgeP12 as 1 for ages below 17 and over 65, and 0 for all other
ages
dep[ , 2:13]  <- lapply(dep[ ,2:13] , FUN = function(x) recode(x, "0:17=1;18:64=0;65:100=1"))

#Add number of dependants
dep$n.dep <- rowSums(dep[,c(2:13)], na.rm = T)

#Find percentage of dependents
dep$p.dep <- dep$n.dep/DHSWI$sumResidents

#Add to SDC dataframe
SDC$DEP <- dep$p.dep
```


---

#1.4 Health Component (HC)
> A. Average of reported household health (HLT)

```{r HC dataframe}
#Create new dataframe for HC variables that includes country and gender
HC <- select(data, Questionnaire, Country, HH.Gender)
```


#1.4.A Average of reported household health (HLT)

```{r HLT - Average of reported household health}
````
````

```
#Calculate average household health status
#Add in sumRes to original dataset
data$sum.res <- DHSWI$sumResidents
#Subset health questions
health <- select(data, Questionnaire, sum.res, Q3HealthstatusP1, Q3HealthstatusP2,
Q3HealthstatusP3, Q3HealthstatusP4, Q3HealthstatusP5, Q3HealthstatusP6,
Q3HealthstatusP7, Q3HealthstatusP8, Q3HealthstatusP9, Q3HealthstatusP10,
Q3HealthstatusP11, Q3HealthstatusP12)

health$total <- rowSums(health[3:14], na.rm=T)
health$avg.health <- (health$total/health$sum.res)
HC$HLT <- health$avg.health
```

#1.5 Social Networks Component (SNC)
> A. Credit use
> B. Govermental support and conservnacy support
> D. Participation in community institutions

```{r SNC dataframe}
#Create new dataframe for SNC variables that includes gender and country
SNC <- select(data, Questionnaire, Country, HH.Gender)
```

#1.5.A Credit use (CRD)

```{r CRD - Credit Use}

#Subset data
cred <- select(data, Questionnaire, Q16CalcTotalHHloansreceivedfromfriendsorfamily,
Q16CalcTotalHHreceivedinformalloans, Q16CalcTotalHHreceivedininformalloans)

#Calculate Credit Use
cred$total <- rowSums(cred[2:4], na.rm=T)

#Assign to SNC dataframe
SNC$CRD <- cred$total/data$ExchangeRatetoUSused
```

#1.5.B Governmental and Conservancy Support (GVT)

```
```{r GVT - Governmental and Conservancy Support}

#Subset data
gvt <- select(data, Questionnaire, Q17CalcTotalamtHHreceivedfromgovtpension,
Q17CalcTotalamtHHreceivedfromgovtOVCsupport,
Q17CalcTotalamtHHreceivedfromgovtreliefprogram,
Q17CalcTotalamtHHreceivedfromconservancytrustcashpmt)

#Calculate Gvt Support
gvt$total <- rowSums(gvt[2:5], na.rm=T)

#Assign to SNC df
SNC$GVT <- gvt$total/data$ExchangeRatetoUSused

#Subtract from household income to reduce redundancy
HWC$INC <- HWC$INC - SNC$GVT
```

#1.5.C Participation in community institutions (PCI)

```{r Participation in community institutions - PCI}
#Subset data related to PCI
PCI <- select(data, Questionnaire, Q46numtimeskgotlaattendedinpastyear,
Q47numtimeschurchattendedinpastyear, Q48TotalnumofinstitutionsHHinvolvedin)

#Recode church and kgotla attendance to 1 (indicating participation) and 0 (indicating no
participation)
PCI[2:3] <- ifelse(PCI[2:3] >= 1, 1, 0)

#Calculate total institutionas participation
PCI$total <- rowSums(PCI[2:4], na.rm=T)

#Assign to SNC df
SNC$PCI <- PCI$total
```

#1.6 Water Component (WC)
> A. Drinking water travel time
> B. Drinking water source

```{r WC dataframe}
```
```

```
#Create new dataframe for WC variables that includes gender and country
WC <- select(data, Questionnaire, Country, HH.Gender)
```

#1.6.A Drinking water travel time (DWT)

```{r Drinking water travel time - DWT}
#Calculate average water travel distance from wet and dry season times

#Subset
water <- select(data, Questionnaire, Q52Drinkingwaterdryseasontraveltime,
Q52Drinkingwaterwetseasontraveltime)

#Calculate
water$avg <- (rowSums(water[2:3], na.rm=T)/2)

#Assign to WC df
WC$DWT <- water$avg
```

#1.6.B Drinking water source (DWS)

```{r Drinking water source - DWS}
#Calculate water source rating
#Rated as 1 - Private, 2 - Communal Closed (Community standpipe, Wellpoint/borehole/Use
neighbors), 3 - Open River
ws <- select (data, Questionnaire)
ws$w.source <- data$Q25Watersource
ws$w.source <- recode(ws$w.source, "4=2")
ws$w.source <- recode(ws$w.source, "5=2")
ws$w.source <- recode(ws$w.source, "6=2")

WC$DWS <- ws$w.source
```

#1.7 Livelihood Diversity Component (LDC)
> A. Livelihood diversification index
> B. Natural resource reliance
> C. Agricultural reliance

```{r LDC dataframe}
```

```
#Create new dataframe for LDC variables that includes gender and country
LDC <- select(data, Questionnaire, Country, HH.Gender)
```

#1.7.A Livelihood diversification index (LDI)

```{r Livelihood Diversification Index - LDI}

#Calculate Livelihood Diversification Index
#Inverse of (Number of livelihood activities +1)

#Livelihood Activities (10): Tourism Permanant, Tourism Casual, Shop/Office Work, Small business of Crafts, Other Sources, Labor Program, Remittances, Farming, Ranching, Natural resource collection

LDI <- select(data, Questionnaire, Q14PermanentincomefromtourismnuminHH,
Q14PermanentincomefromshopofficeworknuminHH,
Q14CasualwageincomefromtourismnuminHH, Q14SmallbusinessesorcraftsnuminHH,
Q14OtherincomesourcesnuminHH, Q17TakepartinlabourprogramYN,
Q15Totalamountreceivedinremittancesinpastyear, Q33DoesHHnormallygrowcropsYN,
Q41HHkeeplivestockYN)
LDI$nr <- totals$totalnrcash

#Recode so that participation in activities is indicated by a 1, no participation is indicated by a 0
LDI[2:11] <- ifelse(LDI[2:11] >= 1, 1, 0)

#Total of livelihood activities per household
LDI$total <- rowSums(LDI[2:11], na.rm=T)

#Create the index
LDI$LDI <- (1/(LDI$total + 1))

#Assign to LDC dataframe
LDC$LDI <- LDI$LDI
```

#1.7.B Natural resource and agricultural reliance (NAR)
Note: This was previously calculated while calculating Overall Annual Income under the Wealth Component (1.2.B)

```{r Natural resource and agricultural reliance - NAR}
#Percentage of total income (including monatary value of all crops and natural resources) that is harvested natural resources
```

```
#Assign to LDC dataframe
LDC$NAR <- totals$nr.ag.rel
```

#1.8 Environmental Shock Component (ESC)
> A. Agricultural loss due to climate and wildlife (AGL)

```{r ESC dataframe}

#Create new dataframe for ESC variables that includes gender and country
ESC <- select(data, Questionnaire, Country, HH.Gender)
```

#1.8.A Agricultural loss due to climate and wildlife (AGL)

```{r AGL.1 Agricultural loss due to climate}

#Subset Variables for household totals of agricultural loss due to climate

#Create dataframe for agricultural loss
AGC <- select(data, Questionnaire)

#Subset and rename variables for each agricultural environmental loss
AGC$maize <- data$Q38Amtmaizelosttoenvproblems
AGC$sorghum <- data$Q38AmtSorghumlosttoenvproblems
AGC$millet <- data$Q38Amtmilletlosttoenvproblems
AGC$beans <- data$Q38AmtBeanspeanutslosttoenvproblems
AGC$pump <- data$Q38AmtPumpkinsmelonslosttoenvproblems
AGC$other <- data$Q38AmtOtherlosttoenvproblems
AGC$cattle <- data$Q43numcattlelosttoenvproblem
AGC$goats <- data$Q43numgoatslosttoenvproblem
AGC$sheep <- data$Q43numsheeplosttoenvproblem
AGC$donkeys <- data$Q43numdonkeyslosttoenvproblem
AGC$horses <- data$Q43numhorseslosttoenvproblem
AGC$poultry <- data$Q43numpoultrylosttoenvproblem
AGC$pigs <- data$Q43numpigslosttoenvproblem

#Convert numbers lost to cash value using house.price dataframe created when calculating
overall average income (INC) in Wealth Component category (1.2.B)
AGC$maize <- AGC$maize * house.price$maize
AGC$sorghum <- AGC$sorghum * house.price$sorghum
```

```
AGC$millet <- AGC$millet * house.price$millet
AGC$beans <- AGC$beans * house.price$beans
AGC$pump <- AGC$pump * house.price$pump
AGC$other <- AGC$other * house.price$pump
AGC$cattle <- AGC$cattle * house.price$cattle
AGC$goats <- AGC$goats * house.price$goats
AGC$sheep <- AGC$sheep * house.price$sheep
AGC$donkeys <- AGC$donkeys * house.price$donkeys
AGC$horses <- AGC$horses * house.price$horses
AGC$poultry <- AGC$poultry * house.price$poultry
AGC$pigs <- AGC$pigs * house.price$pigs

#Add together totals of climate agriculture loss
AGC$total <- rowSums(AGC[2:14], na.rm = T)

```



```{r AGL.2 Agricultural loss due to wildlife}

#Subset variables for household totals of agricultural loss due to wildlife

#Create dataframe for agricultural loss
AGW <- select(data, Questionnaire)

#Subset and rename variables for each agricultural environmental loss
AGW$maize <- data$Q36Amtmaizelosttowildlife
AGW$sorghum <- data$Q36AmtSorghumlosttowildlife
AGW$millet <- data$Q36Amtmilletlosttowildlife
AGW$beans <- data$Q36AmtBeanspeanutslosttowildlife
AGW$pump <- data$Q36AmtPumpkinsmelonslosttowildlife
AGW$other <- data$Q36AmtOtherlosttowildlife
AGW$cattle <- data$Q42numcattlelosttowildlife
AGW$goats <- data$Q42numgoatslosttowildlife
AGW$sheep <- data$Q42numsheeplosttowildlife
AGW$donkeys <- data$Q42numdonkeyslosttowildlife
AGW$horses <- data$Q42numhorseslosttowildlife
AGW$poultry <- data$Q42numpoultrylosttowildlife
AGW$pigs <- data$Q42numpigslosttowildlife

#Convert numbers lost to cash value using house.price dataframe created when calculating
overall average income (INC) in Wealth Component category (1.2.B)
AGW$maize <- AGW$maize * house.price$maize
AGW$sorghum <- AGW$sorghum * house.price$sorghum
```

```
AGW$millet <- AGW$millet * house.price$millet
AGW$beans <- AGW$beans * house.price$beans
AGW$pump <- AGW$pump * house.price$pump
AGW$other <- AGW$other * house.price$pump
AGW$cattle <- AGW$cattle * house.price$cattle
AGW$goats <- AGW$goats * house.price$goats
AGW$sheep <- AGW$sheep * house.price$sheep
AGW$donkeys <- AGW$donkeys * house.price$donkeys
AGW$horses <- AGW$horses * house.price$horses
AGW$poultry <- AGW$poultry * house.price$poultry
AGW$pigs <- AGW$pigs * house.price$pigs

#Add together totals of climate agriculture loss
AGW$total <- rowSums(AGW[2:14], na.rm = T)

#Add to ESC dataframe
ESC$AG <- AGW$total + AGC$total

```

#RQ1 Part 2. Calculate major components of LVI
> 1. Assess each sub-component and log-transform as needed
> 2. Create LVI Dataframe
> 3. Standardize sub-components and calculate major components
> 4. Calculate Livelihood Vulnerability Index
> 5. Prepare household groups for ANOVA tests
> 6. Run T-tests
> 7. Run ANOVA and Tukey HSD tests
> 8. Run T-tests by country
> 9. Run ANVOA and Tukey HSD tests by country
> 10. Data Visualization - Box and Whisker plots (Tukey HSD test results)
> 11. Data Visuzlization - Radar Charts

#2.1 Create LVI dataframe that includes correct form of each sub-component

```{r 2.1 Assess each sub-component and log transform as needed}
hist(FC$FII)
hist(FC$CDI)
hist(HWC$MWI)
```

````
hist(HWC$INC) #badly skewed
HWC$L.INC <- log(HWC$INC +1) #log of household income
hist(HWC$L.INC)
hist(SDC$EDU)
hist(SDC$DEP)
hist(SNC$GVT)
hist(HC$HLT)
hist(SNC$GVT) #Badly skewed
SNC$L.GVT <- log(SNC$GVT +1) #log of GVT
hist(SNC$L.GVT)
hist(SNC$CRD) #Badly skewed
SNC$L.CRD <- log(SNC$CRD +1)
hist(SNC$PCI)
hist(WC$DWT) #badly skewed
WC$L.DWT <- log(WC$DWT +1) #log of DWT
hist(WC$L.DWT)
hist(WC$DWS)
hist(LDC$LDI)
hist(LDC$NAR) #badly skewed
LDC$L.NAR <- log(LDC$NAR +1) #log of NAR
hist(LDC$L.NAR)
hist(ESC$AG)
ESC$L.AG <- log(ESC$AG +1) #log of combined AG
hist(ESC$L.AG)

```




```{r 2.2 Create LVI dataframe}

#Create LVI dataframe

LVI <- select(data, Questionnaire, Country, HH.Gender)
LVI$HH.Gender <- ifelse(LVI$HH.Gender == 1, 0, 1)

#List of variables that contribute to climate vulnerability (Variable assignment)
#For all sub-catagories small numbers indicate good status, large numbers indicate poor status. Inverse is applied to all variables that indicate good status for a high number

#Food Component (2)
#Food Insecurity
LVI$FII <- FC$FII
#Crop Diversity Index
````

```
LVI$CDI <- FC$CDI

#Household Wealth Component (2)
#Material Wealth Index
LVI$MWI <- HWC$MWI
#Overall Annual income
LVI$INC <- (1/(HWC$L.INC + 1))

#Socio-Demographic Component (2)
#Education of household head
LVI$EDU <- (1/(SDC$EDU +1))
#Percentage of dependents
LVI$DEP <- SDC$DEP

#Health Component (1)
#House Health Status
LVI$HLT <- HC$HLT

#Social Network Component (3)
#Note: High credit use, government and conservancy support considered a positive indicator
of greater ability to rely on social networks.
#Credit Use
LVI$CRD <- (1/(SNC$L.CRD +1))
#Participation in Community institutions
LVI$PCI <- (1/(SNC$PCI+1))
#Conservancy and Government Support
LVI$GVT <- (1/SNC$L.GVT+1)

#Water Component (2)
#Drinking water travel time
LVI$DWT <- WC$L.DWT
#Drinking Water source
LVI$DWS <- WC$DWS

#Livelihood Diversity Component (2)
#Natural Resource and Agricultural Reliance
LVI$NAR <- LDC$L.NAR
#Livelihood diversification Index
LVI$LDI <- LDC$LDI

#Environmental Shock Component (1)
#Agriculture loss due to climate and wildlife
LVI$AG <- ESC$L.AG
```

````
```

```{r 2.3 Standardize sub-components and calculate major components}
#Standardization of all Sub-components and Calculation of Major Components

#Equation for standardization of subcomponents is: Standardized sub-component =
(household value - sub component minimum)/(subcomponent maximum - sub component
minimum)
#Equation for calculation of major components is an average of all subcomponents

#Create new dataframes for standardized sub-components and calculated major components
LVI.s <- select(LVI, Questionnaire, HH.Gender, Country)
LVI.m <- select(LVI, Questionnaire, HH.Gender, Country)

#Food Component
#Sub-component standardization
LVI.s$FII <- ((LVI$FII - min(LVI$FII, na.rm= T))/(max(LVI$FII, na.rm = T) - min(LVI$FII, na.rm= T)))
LVI.s$CDI <- ((LVI$CDI - min(LVI$CDI, na.rm= T))/(max(LVI$CDI, na.rm = T) - min(LVI$CDI,
na.rm= T)))
#Major-component calculation
LVI.m$FC <- (rowSums(cbind(LVI.s$FII,LVI.s$CDI),na.rm=TRUE)/2)

#Wealth Component
#Sub-component standardization
LVI.s$MWI <- ((LVI$MWI - min(LVI$MWI, na.rm= T))/(max(LVI$MWI, na.rm = T) -
min(LVI$MWI, na.rm= T)))
LVI.s$INC <- ((LVI$INC - min(LVI$INC, na.rm= T))/(max(LVI$INC, na.rm = T) - min(LVI$INC,
na.rm= T)))
#Major-component calculation
LVI.m$HWC <- (rowSums(cbind(LVI.s$MWI,LVI.s$INC),na.rm=TRUE)/2)

#Socio-Demographic Component
#Sub-component standardization
LVI.s$EDU <- ((LVI$EDU - min(LVI$EDU, na.rm= T))/(max(LVI$EDU, na.rm = T) - min(LVI$EDU,
na.rm= T)))
LVI.s$DEP <- ((LVI$DEP - min(LVI$DEP, na.rm= T))/(max(LVI$DEP, na.rm = T) - min(LVI$DEP,
na.rm= T)))
#Major-component calculation
LVI.m$SDC <- (rowSums(cbind(LVI.s$EDU,LVI.s$DEP),na.rm=TRUE)/2)

#Health Component
#Sub-component standardization
````

```
LVI.s$HLT <- ((LVI$HLT - min(LVI$HLT, na.rm= T))/(max(LVI$HLT, na.rm = T) - min(LVI$HLT,
na.rm= T)))
#Major-component calculation
LVI.m$HC <- LVI.s$HLT

#Social Networks Component
#Sub-component standardization
LVI.s$CRD <- ((LVI$CRD - min(LVI$CRD, na.rm= T))/(max(LVI$CRD, na.rm = T) - min(LVI$CRD,
na.rm= T)))
LVI.s$GVT <- ((LVI$GVT - min(LVI$GVT, na.rm= T))/(max(LVI$GVT, na.rm = T) - min(LVI$GVT,
na.rm= T)))
LVI.s$PCI <- ((LVI$PCI - min(LVI$PCI, na.rm= T))/(max(LVI$PCI, na.rm = T) - min(LVI$PCI, na.rm=
T)))
#Major Component calculation
LVI.m$SNC <- (rowSums(cbind(LVI.s$CRD,LVI.s$GVT,LVI.s$PCI),na.rm=TRUE)/3)

#Water Component
#Sub-component standardization
LVI.s$DWT <- ((LVI$DWT - min(LVI$DWT, na.rm= T))/(max(LVI$DWT, na.rm = T) -
min(LVI$DWT, na.rm= T)))
#LVI.s$LWT <- ((LVI$LWT - min(LVI$LWT, na.rm= T))/(max(LVI$LWT, na.rm = T) - min(LVI$LWT,
na.rm= T)))
LVI.s$DWS <- ((LVI$DWS - min(LVI$DWS, na.rm= T))/(max(LVI$DWS, na.rm = T) -
min(LVI$DWS, na.rm= T)))
#Major Component Calculation
LVI.m$WC <- (rowSums(cbind(LVI.s$DWT,LVI.s$DWS),na.rm=TRUE)/2)

#Livelihood Diversity Component
LVI.s$NAR <- ((LVI$NAR - min(LVI$NAR, na.rm= T))/(max(LVI$NAR, na.rm = T) - min(LVI$NAR,
na.rm= T)))
LVI.s$LDI<- ((LVI$LDI - min(LVI$LDI, na.rm= T))/(max(LVI$LDI, na.rm = T) - min(LVI$LDI, na.rm=
T)))
#Major Component Calculation
LVI.m$LDC <- (rowSums(cbind(LVI.s$NAR,LVI.s$LDI),na.rm=TRUE)/2)

#Environmental Shock Component
LVI.s$AG <- ((LVI$AG - min(LVI$AG, na.rm= T))/(max(LVI$AG, na.rm = T) - min(LVI$AG, na.rm=
T)))
#Major Component Calculation
LVI.m$ESC <- LVI.s$AG
```

````
```{r 2.4 Calculate Livelihood Vulnerability Index}
#Calculate Livelihood Vulnerability Index (LVI)
````
````

```
#Calculated by multiplying each major component score by (number of subcomponents * sum of sub component weightings) to acheive a weighted major component. All weighted components are added together and divided by (the total number of subcomponents * sum of sub component weightings).
#Formula: ((nSC*wSC)(MC) + (nSC*wSC)(MC) + (nSC*wSC)(MC)...)/(nSC*wSC)+(nSC*wSC)+(nSC*wSC)...

#Weights (equal to number of subcomponents in each category):
wFC <- 2
wHWC <- 2
wSDC <- 2
wHC <-  1
wSNC <- 3
wWC <- 2
wLDC <- 2
wESC <- 1


LVI.m$LVI <- ((LVI.m$FC * wFC) + (LVI.m$HWC * wHWC) + (LVI.m$SDC * wSDC) + (LVI.m$HC * wHC) + (LVI.m$SNC * wSNC) + (LVI.m$WC * wWC) + (LVI.m$LDC * wLDC) + (LVI.m$ESC * wESC))/(wFC + wHWC + wSDC + wHC + wSNC + wWC + wESC)

LVI$LVI <- LVI.m$LVI
```

```{r 2.5 Prepare household groups for ANOVA tests}
#Subset information related to gender and status of hosuehold memebers
gender <- select(data, Questionnaire, HH.Gender, Q3RelationtoHHheadP1, Q3RelationtoHHheadP2, Q3RelationtoHHheadP3, Q3RelationtoHHheadP4, Q3RelationtoHHheadP5, Q3RelationtoHHheadP6, Q3RelationtoHHheadP7, Q3RelationtoHHheadP8, Q3RelationtoHHheadP9, Q3RelationtoHHheadP10, Q3RelationtoHHheadP11, Q3RelationtoHHheadP12)

#Change all residents who reported as spouse to a 1, change all other values to 0
gender[3:14] <- ifelse(gender[3:14] == 2, 1, 0)
#Add together to determine how many spouses in each household
gender$spouse <- rowSums(gender[3:14], na.rm=T)

#Subset into female de jure hosueholds
f.dj <- subset(gender, HH.Gender == 3)
#Subset into female de facto hosueholds
f.df <- subset(gender, HH.Gender ==2)
#Subset into male households
```

```
male <- subset(gender, HH.Gender ==1)
#Characterize if male households have a spouse (1) or don't have a spouse (0)
male$spouse <- ifelse(male$spouse == 0,0, 1)
#Subset into male spouse households
m.s <- subset(male, spouse ==1)
#Subset into male no spouse households
m.ns <- subset(male, spouse ==0)

#Label each household type
f.dj$GEN <- "FDJ"
f.df$GEN <- "FDF"
m.s$GEN <- "MSP"
m.ns$GEN <- "MNS"

#Merge together all household types
gend2 <- rbind(f.dj, f.df, m.s, m.ns)
#Arrange df according to questionnaire
gend2 <- gend2 %>% arrange(Questionnaire)
```

```{r 2.6 Run T-tests}
#Results for t-tests run on the overall LVI and each major component by gender of household head

t.test(LVI.m$LVI ~LVI.m$HH.Gender)
t.test(LVI.m$FC ~ LVI.m$HH.Gender)
t.test(LVI.m$HWC ~ LVI.m$HH.Gender)
t.test(LVI.m$SDC ~ LVI.m$HH.Gender)
t.test(LVI.m$HC ~ LVI.m$HH.Gender)
t.test(LVI.m$SNC ~ LVI.m$HH.Gender)
t.test(LVI.m$WC ~ LVI.m$HH.Gender)
t.test(LVI.m$LDC ~ LVI.m$HH.Gender)
t.test(LVI.m$ESC ~ LVI.m$HH.Gender)

```

```{r 2.7 Run ANOVA and Tukey HSD tests}
#ANOVA Test by new gender groups

#Add gendered groups to LVI.m dataframe
LVI.m$GEN <- gend2$GEN
```

```
#Overall LVI
aov.g <- aov(LVI ~ GEN, data = LVI.m) #run anova
summary(aov.g) #summary of ANOVA
TukeyHSD(aov.g) #run Tukey HSD to find which groups significance is between
#Food MC
FC.an <- aov(FC ~ GEN, data = LVI.m)
summary(FC.an)
TukeyHSD(FC.an)
#Household Wealth MC
HWC.an <- aov(HWC ~ GEN, data = LVI.m)
summary(HWC.an)
TukeyHSD(HWC.an)
#Socio-Demographic MC
SDC.an <- aov(SDC ~ GEN, data = LVI.m)
summary(SDC.an)
TukeyHSD(SDC.an)
#Health MC
HC.an <- aov(HC ~ GEN, data = LVI.m)
summary(HC.an)
TukeyHSD(HC.an)
#Social Networks MC
SNC.an <- aov(SNC ~ GEN, data = LVI.m)
summary(SNC.an)
TukeyHSD(SNC.an)
#Livelihood Diversity MC
LDC.an <- aov(LDC ~ GEN, data = LVI.m)
summary(LDC.an)
TukeyHSD(LDC.an)
#Water MC
WC.an <- aov(WC ~ GEN, data = LVI.m)
summary(WC.an)
TukeyHSD(WC.an)
#Environmental Shock MC
ESC.an <- aov(ESC ~ GEN, data = LVI.m)
summary(ESC.an)
TukeyHSD(ESC.an)
'''

'''{r 2.8 Run T-test by country}
#Subset Botswana
B.LVI <- subset(LVI.m, Country == "1")

#Run t-tests on LVI and major components for Botswana
t.test(B.LVI$LVI ~ B.LVI$HH.Gender)
```

```
t.test(B.LVI$FC ~ B.LVI$HH.Gender)
t.test(B.LVI$HWC ~ B.LVI$HH.Gender)
t.test(B.LVI$SDC ~ B.LVI$HH.Gender)
t.test(B.LVI$HC ~ B.LVI$HH.Gender)
t.test(B.LVI$SNC ~ B.LVI$HH.Gender)
t.test(B.LVI$WC ~ B.LVI$HH.Gender)
t.test(B.LVI$LDC ~ B.LVI$HH.Gender)
t.test(B.LVI$ESC ~ B.LVI$HH.Gender)

#Subset Namibia
N.LVI <- subset(LVI.m, Country == "2")

#Run t-tests on LVI and major components for Namibia
t.test(N.LVI$LVI ~ N.LVI$HH.Gender)
t.test(N.LVI$FC ~ N.LVI$HH.Gender)
t.test(N.LVI$HWC ~ N.LVI$HH.Gender)
t.test(N.LVI$SDC ~ N.LVI$HH.Gender)
t.test(N.LVI$HC ~ N.LVI$HH.Gender)
t.test(N.LVI$SNC ~ N.LVI$HH.Gender)
t.test(N.LVI$WC ~ N.LVI$HH.Gender)
t.test(N.LVI$LDC ~ N.LVI$HH.Gender)
t.test(N.LVI$ESC ~ N.LVI$HH.Gender)

#Subset Zambia
Z.LVI <- subset(LVI.m, Country == "3")

#Run t-tests on LVI and major components for Zambia
t.test(Z.LVI$LVI ~ Z.LVI$HH.Gender)
t.test(Z.LVI$FC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$HWC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$SDC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$HC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$SNC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$WC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$LDC ~ Z.LVI$HH.Gender)
t.test(Z.LVI$ESC ~ Z.LVI$HH.Gender)
` ` `

` ` `{r 2.9 Run ANOVA and Tukey HSD by Country}

#Botswana
aov.B <- aov(LVI ~ GEN, data = B.LVI) #Run ANOVA for Botswana LVI (step 1)
summary(aov.B) #Summary of Botswana LVI ANOVA (step 2)
TukeyHSD(aov.B) #Tukey HSD for Botswana LVI (step 3)
```

```
FC.b <- aov(FC ~ GEN, data = B.LVI) #Repeat 3 above steps for all major componenets
summary(FC.b)
TukeyHSD(FC.b)
HWC.b <- aov(HWC ~ GEN, data = B.LVI)
summary(HWC.b)
TukeyHSD(HWC.b)
SDC.b <- aov(SDC ~ GEN, data = B.LVI)
summary(SDC.b)
TukeyHSD(SDC.b)
HC.b <- aov(HC ~ GEN, data = B.LVI)
summary(HC.b)
TukeyHSD(HC.b)
SNC.b <- aov(SNC ~ GEN, data = B.LVI)
summary(SNC.b)
TukeyHSD(SNC.b)
LDC.b <- aov(LDC ~ GEN, data = B.LVI)
summary(LDC.b)
TukeyHSD(LDC.b)
WC.b <- aov(WC ~ GEN, data = B.LVI)
summary(WC.b)
TukeyHSD(WC.b)
ESC.b <- aov(ESC ~ GEN, data = B.LVI)
summary(ESC.b)
TukeyHSD(ESC.b)

#Nambia
aov.n <- aov(LVI ~ GEN, data = N.LVI) #Repeat above Botswana steps for Namibia
summary(aov.n)
TukeyHSD(aov.n)
FC.n <- aov(FC ~ GEN, data = N.LVI)
summary(FC.n)
TukeyHSD(FC.n)
HWC.n <- aov(HWC ~ GEN, data = N.LVI)
summary(HWC.n)
TukeyHSD(HWC.n)
SDC.n <- aov(SDC ~ GEN, data = N.LVI)
summary(SDC.n)
TukeyHSD(SDC.n)
HC.n <- aov(HC ~ GEN, data = N.LVI)
summary(HC.n)
TukeyHSD(HC.n)
SNC.n <- aov(SNC ~ GEN, data = N.LVI)
summary(SNC.n)
TukeyHSD(SNC.n)
```

```
LDC.n <- aov(LDC ~ GEN, data = N.LVI)
summary(LDC.n)
TukeyHSD(LDC.n)
WC.n <- aov(WC ~ GEN, data = N.LVI)
summary(WC.n)
TukeyHSD(WC.n)
ESC.n <- aov(ESC ~ GEN, data = N.LVI)
summary(ESC.n)
TukeyHSD(ESC.n)

#Zambia
aov.z <- aov(LVI ~ GEN, data = Z.LVI) #Repeat above botswana steps for Zambia
summary(aov.z)
TukeyHSD(aov.z)
FC.z <- aov(FC ~ GEN, data = Z.LVI)
summary(FC.z)
TukeyHSD(FC.z)
HWC.z <- aov(HWC ~ GEN, data = Z.LVI)
summary(HWC.z)
TukeyHSD(HWC.z)
SDC.z <- aov(SDC ~ GEN, data = Z.LVI)
summary(SDC.z)
TukeyHSD(SDC.z)
HC.z <- aov(HC ~ GEN, data = Z.LVI)
summary(HC.z)
TukeyHSD(HC.z)
SNC.z <- aov(SNC ~ GEN, data = Z.LVI)
summary(SNC.z)
TukeyHSD(SNC.z)
LDC.z <- aov(LDC ~ GEN, data = Z.LVI)
summary(LDC.z)
TukeyHSD(LDC.z)
WC.z <- aov(WC ~ GEN, data = Z.LVI)
summary(WC.z)
TukeyHSD(WC.z)
ESC.z <- aov(ESC ~ GEN, data = Z.LVI)
summary(ESC.z)
TukeyHSD(ESC.z)
```

```{r 2.10 Data visualization - Box and Whisker Plots (Tukey HSD test results)}
```

```
#Label different sets of groups to "star" indicating significance. Information from previously run Tukey HSD tests
star1 <- list(c("FDJ", "FDF"), c("MSP", "FDJ"), c("MNS", "MSP"),  c("FDF", "MNS"))
star2 <- list(c("FDJ", "MSP"))
star3 <- list(c("FDJ", "MSP"), c("MSP", "MNS"))

#Box and Whisker plot for overall LVI by gendered households
ggboxplot(LVI.m, x = "GEN", y = "LVI") +
stat_compare_means(comparisons = star1, label = "p.signif")
#B&W plot for FC by gendered households
ggboxplot(LVI.m, x = "GEN", y = "FC") + stat_compare_means(method = "anova", label.y = 1.6)
+
stat_compare_means(comparisons = star1, label = "p.signif") + theme_minimal() + labs(x =
"Household Head", y = "Food Component Score")
#B&W plot for SDC by gendered households
ggboxplot(LVI.m, x = "GEN", y = "SDC") + stat_compare_means(method = "anova", label.y =
1.1) +
stat_compare_means(comparisons = star2, label = "p.signif") + theme_minimal() + labs(x =
"Household Head", y = "SD Component Score")
#B&W plot for SNC by gendered households
ggboxplot(LVI.m, x = "GEN", y = "SNC") + stat_compare_means(method = "anova", label.y =
1.2) +
stat_compare_means(comparisons = star2, label = "p.signif") + theme_minimal() + labs(x =
"Household Head", y = "SN Component Score")
#B&W plot for LDC by gendered households
ggboxplot(LVI.m, x = "GEN", y = "LDC") + stat_compare_means(method = "anova", label.y =
1.4) +
stat_compare_means(comparisons = star3, label = "p.signif") + theme_minimal() + labs(x =
"Household Head", y = "LD Component Score")
```

```{r 2.11 Data Visualization - Radar charts}
#Radar Graphs

#Create vectors for later chart labels
gen <- c("M", "F") #Gender of HH head vector
cty <- c("Botswana", "Namibia", "Zambia") #country vector
hh <- c("FDF", "FDJ", "MNS", "MSP") #household head vector

#Create vectors for colors
colors_border=c( rgb(0.2,0.5,0.5,0.9), rgb(0.8,0.2,0.5,0.9) , rgb(0.7,0.5,0.1,0.9) )
colors_in=c( rgb(0.2,0.5,0.5,0.4), rgb(0.8,0.2,0.5,0.4) , rgb(0.7,0.5,0.1,0.4) )
cHH <- c("indianred1", "chartreuse4", "turquoise3", "mediumpurple")
```

```
#Total LVI dataset Radar chart
LVI.r <- select(LVI.m, c(2, 4:11)) #select relevant info
LVI.r$group <- "1" #assign dummy variable for aggregation
LVI.r.a <- aggregate(LVI.r[, 2:9], list(LVI.r$group), mean, na.rm = T) #aggregate info
LVI.r.a$group <- "NULL" #delete dummy variable
LVI.2 <- select(LVI.r.a, c(2:9)) #select only major components
LVI.2 <- rbind(rep(.8,8) , rep(0,8) , LVI.2) #add information about radar chart limits
radarchart(LVI.2, axistype = 1, plwd = 4, plty = 1, cglcol = "grey", cglty = 1, axislabcol = "grey",
caxislabels = seq(0,1,.2), cglwd = 0.8, vlcex = 0.8, title=paste("LVI Major Components"))
#create the chart

#Total LVI displayed by gender
LVI.g <- select(LVI.m, c(2, 4:11))
LVI.g.a <- aggregate(LVI.g[, 2:9], list(LVI.g$HH.Gender), mean, na.rm = T)
LVI.g.2 <- select(LVI.g.a, c(2:9))
LVI.g.2 <- rbind(rep(.8,8) , rep(0,8) , LVI.g.2)
radarchart(LVI.g.2, axistype = 1, pcol = colors_border, plwd = 4, plty = 1, cglcol = "grey", cglty =
1, axislabcol = "grey", caxislabels = seq(0,1,.2), cglwd = 0.8, vlcex = 0.8, title=paste("LVI Major
Components by Gender"))
legend(x=1.5, y=1, legend = gen, bty = "n", pch=20 , col=colors_border, cex=1.2, pt.cex=3)

#Total LVI displayed by gendered household groups
LVI.hh <- select(LVI.m, c(4:11, 13))
LVI.hh.a <- aggregate(LVI.hh[, 1:8], list(LVI.hh$GEN), mean, na.rm = T)
LVI.hh.a <- select(LVI.hh.a, c(2:9))
LVI.hh.a <- rbind(rep(.8,8) , rep(0,8) , LVI.hh.a)
radarchart(LVI.hh.a, axistype = 1, pcol = cHH, plwd = 4, plty = 1, cglcol = "grey", cglty = 1,
axislabcol = "grey", caxislabels = seq(0,1,.2), cglwd = 0.8, vlcex = 0.8, title=paste("LVI Major
Components by Gendered Household Groups"))
legend(x=1.5, y=1, legend = hh, bty = "n", pch=20 , col=cHH, cex=1.2, pt.cex=3)

#Total LVI displayed by Country
LVI.c <- select(LVI.m, c(3:11))
LVI.c.a <- aggregate(LVI.c[, 2:9], list(LVI.c$Country), mean, na.rm = T)
LVI.c.2 <- select(LVI.c.a, c(2:9))
LVI.c.2 <- rbind(rep(.8,8) , rep(0,8) , LVI.c.2)
radarchart(LVI.c.2, axistype = 1, pcol = colors_border, plwd = 4, plty = 1, cglcol="grey", cglty =
1, axislabcol = "grey", caxislabels = seq(0,1,.2), cglwd=0.8, vlcex=0.8, title=paste("LVI Major
Components by Country"))
legend(x=1.3, y=1, legend = cty, bty = "n", pch=15 , col=colors_border, cex=1.2, pt.cex=3)

#Botswana LVI displayed by Gender
B.LVI.a <- select(B.LVI, c(2, 4:12))
```

```
B.LVI.a <- aggregate(B.LVI.a[,2:10], list(B.LVI.a$HH.Gender), mean, na.rm = T)
B.LVI.a <- select(B.LVI.a, c(2:9))
B.LVI.a <- rbind(rep(.8,8) , rep(0,8), B.LVI.a)
radarchart(B.LVI.a, axistype = 1, pcol = colors_border, plwd = 4, plty = 1, cglcol="grey", cglty =
1, axislabcol = "grey", caxislabels = seq(0,1,.2), cglwd=0.8, vlcex=0.8, title=paste("Botswana LVI
Major Components by Gender"))
legend(x=1.5, y=1, legend = gen, bty = "n", pch=20 , col=colors_border, cex=1.2, pt.cex=3)

#Zambia LVI displayed by Gender
Z.LVI.a <- select(Z.LVI, c(2, 4:12))
Z.LVI.a <- aggregate(Z.LVI.a[,2:10], list(Z.LVI.a$HH.Gender), mean, na.rm = T)
Z.LVI.a <- select(Z.LVI.a, c(2:9))
Z.LVI.a <- rbind(rep(.8,8) , rep(0,8), Z.LVI.a)
radarchart(Z.LVI.a, axistype = 1, pcol = colors_border, plwd = 4, plty = 1, cglcol="grey", cglty =
1, axislabcol = "grey", caxislabels = seq(0,1,.2), cglwd=0.8, vlcex=0.8, title=paste("Zambia LVI
Major Components by Gender"))
legend(x=1.5, y=1, legend = gen, bty = "n", pch=20 , col=colors_border, cex=1.2, pt.cex=3)

#Namibia LVI displayed by Gender
N.LVI.a <- select(N.LVI, c(2, 4:12))
N.LVI.a <- aggregate(N.LVI.a[,2:10], list(N.LVI.a$HH.Gender), mean, na.rm = T)
N.LVI.a <- select(N.LVI.a, c(2:9))
N.LVI.a <- rbind(rep(.8,8) , rep(0,8), N.LVI.a)
radarchart(N.LVI.a, axistype = 1, pcol = colors_border, plwd = 4, plty = 1, cglcol="grey", cglty =
1, axislabcol = "grey", caxislabels = seq(0,1, .2), cglwd=1, vlcex=1, title=paste("Namibia LVI
Major Components by Gender"))
legend(x=1.5, y=1, legend = gen, bty = "n", pch=20 , col=colors_border, cex=1.2, pt.cex=3)

#Botswana LVI by household head groups
B.LVI.b <- select(B.LVI, c(4:11, 13))
B.LVI.b <- aggregate(B.LVI.b[,1:8], list(B.LVI.b$GEN), mean, na.rm = T)
B.LVI.b <- select(B.LVI.b, c(2:9))
B.LVI.b <- rbind(rep(.8,8) , rep(0,8), B.LVI.b)
radarchart(B.LVI.b, axistype = 1, pcol = cHH, plwd = 4, plty = 1, cglcol="grey", cglty = 1,
axislabcol = "grey", caxislabels = seq(0,1, .2), cglwd=1, vlcex=1)
legend(x=1.5, y=1, legend = hh, bty = "n", pch=20 , col=cHH, cex=1.2, pt.cex=3)

#Namibia LVI by household head groups
N.LVI.b <- select(N.LVI, c(4:11, 13))
N.LVI.b <- aggregate(N.LVI.b[,1:8], list(N.LVI.b$GEN), mean, na.rm = T)
N.LVI.b <- select(N.LVI.b, c(2:9))
N.LVI.b <- rbind(rep(1,8) , rep(0,8), N.LVI.b)
radarchart(N.LVI.b, axistype = 1, pcol = cHH, plwd = 4, plty = 1, cglcol="grey", cglty = 1,
axislabcol = "grey", caxislabels = seq(0,1, .2), cglwd=1, vlcex=1)
```

```
legend(x=1.5, y=1, legend = hh, bty = "n", pch=20 , col=cHH, cex=1.2, pt.cex=3)

#Zambia LVI by household head groups
Z.LVI.b <- select(Z.LVI, c(4:11, 13))
Z.LVI.b <- aggregate(Z.LVI.b[,1:8], list(Z.LVI.b$GEN), mean, na.rm = T)
Z.LVI.b <- select(Z.LVI.b, c(2:9))
Z.LVI.b <- rbind(rep(.8,8) , rep(0,8), Z.LVI.b)
radarchart(Z.LVI.b, axistype = 1, pcol = cHH, plwd = 4, plty = 1, cglcol="grey", cglty = 1,
axislabcol = "grey", caxislabels = seq(0,1, .2), cglwd=1, vlcex=1)
legend(x=1.5, y=1, legend = hh, bty = "n", pch=20 , col=cHH, cex=1.2, pt.cex=3)
```

#RQ2. Characterize CBC influence through employment and participation
> Part 1: Characterize employment in study group
> Part 2: Other tests related to CBC participation and nvolvement
---

#Part 1: Characterize employment in study group
> 1. Subset data
> 2. Label job titles and gender
> 3. Assign job titles to catagories
> 4. Find counts by gender for entire dataset
> 5. Find counts by gender type of household
> 6. Employment data visualization

```{r 2.1.1 Subset data}

data$GEN <- LVI.m$GEN
P1 <- select(data, Questionnaire, VillageArea, Country, Q1GenderofHHhead, Q3JobTitleP1,
GEN)
P2 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP2, Q3JobTitleP2, GEN)
P3 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP3, Q3JobTitleP3, GEN)
P4 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP4, Q3JobTitleP4, GEN)
P5 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP5, Q3JobTitleP5, GEN)
P6 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP6, Q3JobTitleP6, GEN)
P7 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP7, Q3JobTitleP7, GEN)
P8 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP8, Q3JobTitleP8, GEN)
P9 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP9, Q3JobTitleP9, GEN)
P10 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP10, Q3JobTitleP10, GEN)
P11 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP11, Q3JobTitleP11, GEN)
P12 <- select(data, Questionnaire, VillageArea, Country, Q3GenderP12, Q3JobTitleP12, GEN)
```

```
P1 <- rename(P1, Gender = Q1GenderofHHhead, Job = Q3JobTitleP1)
P2 <- rename(P2, Gender = Q3GenderP2, Job = Q3JobTitleP2)
P3 <- rename(P3, Gender = Q3GenderP3, Job = Q3JobTitleP3)
P4 <- rename(P4, Gender = Q3GenderP4, Job = Q3JobTitleP4)
P5 <- rename(P5, Gender = Q3GenderP5, Job = Q3JobTitleP5)
P6 <- rename(P6, Gender = Q3GenderP6, Job = Q3JobTitleP6)
P7 <- rename(P7, Gender = Q3GenderP7, Job = Q3JobTitleP7)
P8 <- rename(P8, Gender = Q3GenderP8, Job = Q3JobTitleP8)
P9 <- rename(P9, Gender = Q3GenderP9, Job = Q3JobTitleP9)
P10 <- rename(P10, Gender = Q3GenderP10, Job = Q3JobTitleP10)
P11 <- rename(P11, Gender = Q3GenderP11, Job = Q3JobTitleP11)
P12 <- rename(P12, Gender = Q3GenderP12, Job = Q3JobTitleP12)

#Add in participant ID to each subset
P1$P <- "P1"
P2$P <- "P2"
P3$P <- "P3"
P4$P <- "P4"
P5$P <- "P5"
P6$P <- "P6"
P7$P <- "P7"
P8$P <- "P8"
P9$P <- "P9"
P10$P <- "P10"
P11$P <- "P11"
P12$P <- "P12"

#Join all participant subsets into a single table
emp <- rbind(P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12)

#Recode both female gender catagories as 2
emp$Gender <- recode(emp$Gender, "3=2")
` ` `

` ` `{r 2.1.2 Label job titles and gender}

#Re-label gender to M and W
emp$Gender[emp$Gender ==1] <- "M"
emp$Gender[emp$Gender == 2] <- "W"

#Create new variable for job title
emp$title <- emp$Job
```

```
#Label each job title based on code
emp$title[emp$title==1] <- "Herder"
emp$title[emp$title==2] <- "Safari Guide"
emp$title[emp$title==3] <- "Driver"
emp$title[emp$title==4] <- "National Service"
emp$title[emp$title==5] <- "Small business owner"
emp$title[emp$title==6] <- "Lodge Employee"
emp$title[emp$title==7] <- "Teacher"
emp$title[emp$title==8] <- "Casual Labourer"
emp$title[emp$title==9] <- "Builder"
emp$title[emp$title==10] <- "Clerk"
emp$title[emp$title==11] <- "Chef"
emp$title[emp$title==12] <- "Cleaner"
emp$title[emp$title==13] <- "Nurse/health"
emp$title[emp$title==14] <- "Security guard"
emp$title[emp$title==15] <- "Govt. Tech. Officer"
emp$title[emp$title==16] <- "Police"
emp$title[emp$title==17] <- "Headman"
emp$title[emp$title==18] <- "Mechanic"
emp$title[emp$title==19] <- "Manager"
emp$title[emp$title==20] <- "Soldier"
emp$title[emp$title==21] <- "Miner"
emp$title[emp$title==22] <- "Other"
emp$title[emp$title==23] <- "Escort guide/Conservancy Rep"
emp$title[emp$title==24] <- "Unknown"
emp$title[emp$title==25] <- "Unknown"
emp$title[emp$title==33] <- "Unknown"
```

```{r 2.1.3 Assign job titles to categories}
#Assign each job title to a category (directly generated by conservancy or  government, related to conservancy, traditional/not related to consv) held by each gender

#Cons <- Safari Guid, National service, Escort Guide, Govt. Tech. Off
#Cons.R <- Driver, Small Business, Lodge Employee, Clerk, Chef, Cleaner, Security Guard, Mechanic, Manager
#Trad <- Herder, Teacher, Casual Lab, Builder, Nurse/health, Police, Headman, Soldier, Miner
#Unknown <- 22 (Other), 24, 25, 33 (Unknown)

#Create new variable for category
emp$cat <- emp$Job

#Code each job-holder into a category
emp$cat<-recode(emp$cat,"c(2,4,15,23)='Cons'")
```

````
emp$cat<-recode(emp$cat, "c(3, 5, 6, 10, 11, 12, 14, 18, 19)='Cons.R'")
emp$cat<-recode(emp$cat, "c(1, 7, 8, 9, 13, 16, 17, 20, 21)='Trad'")
emp$cat<-recode(emp$cat, "c(22, 24, 25, 33)= 'Unkn'")
```

```{r 2.1.4 Find counts by gender for entire dataset}
emp$n <- 1 #apply 1 to each record for counts
#Aggregate job type by gender
emp.job <- aggregate(emp$n, by=list(Gender=emp$Gender, Job=emp$title), FUN = sum)
#Aggregate job category by gender
emp.cat <- aggregate(emp$n, by=list(Gender=emp$Gender, cat=emp$cat), FUN = sum)

#Find total jobs by gender
emp.t <- aggregate(emp.job$x, by=list(Gender=emp.job$Gender), FUN=sum)
```

```{r 2.1.5 Find counts by gender type of household}

#Aggregate job type by gender
emp.job2 <- aggregate(emp$n, by=list(Gender=emp$GEN, Job=emp$title), FUN = sum)
#Aggregate job category by gender
emp.cat2 <- aggregate(emp$n, by=list(Gender=emp$GEN, cat=emp$cat), FUN = sum)

#Find total jobs by gender
emp.t2 <- aggregate(emp.job2$x, by=list(Gender=emp.job2$Gender), FUN=sum)
```

```{r 2.1.6 Employment data visualization}
#Number of participants in each category by gender
ggplot(data=emp.cat, aes(x=cat, y=x, fill=Gender)) +
geom_bar(stat="identity", position=position_dodge()) + theme_minimal()

#Total number of job-holding participants by gender
ggplot(data=emp.t2, aes(x=Gender, y=x, fill = Gender))+
geom_bar(stat="identity", position=position_dodge()) + theme_minimal()

```
---

#Part 2: Other tests related to CBC participation and involvement
> 1. Subset data related to participation in conservancy
> 2. Find gendered percentages of each type of participation
> 3. Data visualization for gendered participation percentages
> 4. Test for gendered significant differences in overall conservancy participation
````
````

```
> 5. Data visualization for overall conservancy participation
> 6. Assess correlations between participation score and LVI scores
> 7. Test for differences in aid received between household types

```{r 2.2.1 Subset data related to participation in conservancy}
#Add transformed data back into data
data$HH.Gender2 <- LVI.m$HH.Gender
#Subset data related to PCN
PCN <- select(data, Questionnaire, Q48.1HeardofConservancyTrustYN,
Q48.1AttendingmeetingsofConservancyTrustYN, Q48.1MemberofConservancyTrustYN,
HH.Gender2)

#Recode variables to 1 (indicating participation) and 0 (indicating no participation)
PCN[2:4] <- ifelse(PCN[2:4] == 1, 1, 0)
PCN[3] <- ifelse(PCN[3] == 1, 2, 0)
PCN[4] <- ifelse(PCN[4] == 1, 3, 0)
#Calculate total institutionas participation
PCN$total <- rowSums(PCN[2:4], na.rm=T)

#Assign to SNC df
PCN$PCN <- PCN$total
```



```{r 2.2.2 Find gendered percentages of each type of participation}

#Create new subset of each type of participation (heard of conservancy, attend meetings,
member in board of trust)
cpart <- select(data, Questionnaire, HH.Gender)
cpart$h <- data$Q48.1HeardofConservancyTrustYN
cpart$a <- data$Q48.1AttendingmeetingsofConservancyTrustYN
cpart$m <- data$Q48AnyHHmemberinBoardofTrustConservancy

#Recode gender
cpart$HH.Gender[cpart$HH.Gender==1] <- "M"
cpart$HH.Gender[cpart$HH.Gender==2|cpart$HH.Gender==3] <- "W"

#Recode yes and no answers
cpart$heard[cpart$h==1] <- "Y"
cpart$heard[cpart$h==2] <- "N"
cpart$attend[cpart$a==1] <- "Y"
cpart$attend[cpart$a==2] <- "N"
cpart$member[cpart$m==1] <- "Y"
cpart$member[cpart$m==2] <- "N"
```
```

```
#Apply 1 to each record in order to perform counts
cpart$n <- 1

#Counts based on gender of each category
cpart.heard <- aggregate(cpart$n, by=list(Gender=cpart$HH.Gender, cpart$heard), FUN =
sum)
cpart.attend <- aggregate(cpart$n, by=list(Gender=cpart$HH.Gender, cpart$attend), FUN =
sum)
cpart.member <- aggregate(cpart$n, by=list(Gender=cpart$HH.Gender, cpart$member), FUN
= sum)
```

```{r 2.2.3 Data visualization for gendered participation percentages}
#Data visualization of heard, attend, member
ggplot(data=cpart.heard, aes(x=Group.2, y=x, fill=Gender)) +
geom_bar(stat="identity", position=position_dodge()) + theme_minimal() +ggtitle("Heard of
Conservancy")

ggplot(data=cpart.attend, aes(x=Group.2, y=x, fill=Gender)) +
geom_bar(stat="identity", position=position_dodge()) + theme_minimal() + ggtitle("Attend
Conservancy Meetings")

ggplot(data=cpart.member, aes(x=Group.2, y=x, fill=Gender)) +
geom_bar(stat="identity", position=position_dodge()) + theme_minimal() + ggtitle("Member
of Conservancy")
```

```{r 2.2.4 Test for gendered significant differences in overall conservancy participation}
#Test for significant differences in participation between gender

#Overall conservancy score
t.test(PCN$PCN ~ PCN$HH.Gender2)
```

```{r 2.2.5 Data visualization for overall conservancy participation}
gencol <- c("indianred1", "turquoise3")
PCN$HH.Gender2 <- as.character(PCN$HH.Gender2)
PCN$HH.Gender2 <- recode(PCN$HH.Gender2, "'0' = 'MHH'; '1' = 'FHH'")
ggplot(PCN, aes(HH.Gender2, PCN, fill = HH.Gender2)) + geom_violin() +
  labs(title="Household Conservancy Participation",
      subtitle = "By Gender of Household Head", caption = "Black point indicates group mean",
      x="Gender of Household Head",
```

```
    y="Conservancy Participation Score") + theme_minimal() + stat_summary(name = "Mean",
fun.y=mean, geom="point", size=2)
```

```{r 2.2.6 Assess correlations between participation score and LVI scores}

cor.test(SNC$PCN, LVI.m$LVI)
cor.test(SNC$PCN, LVI.m$FC)
cor.test(SNC$PCN, LVI.m$HWC)
cor.test(SNC$PCN, LVI.m$HC)
cor.test(SNC$PCN, LVI.m$LDC)
cor.test(SNC$PCN, LVI.m$ESC)
cor.test(SNC$PCN, LVI.m$SDC)
cor.test(SNC$PCN, LVI.m$SNC)
cor.test(SNC$PCN, LVI.m$WC)
```

```{r 2.2.7 Test for differences in aid received by household types}
#T.test bewteen aid received

#Create aid dataframe
aid <- select(LVI.m, HH.Gender, GEN)
aid$aid <- SNC$GVT

#T.test between aid received
t.test(aid$aid ~ aid$HH.Gender)

#Anova test between aid received
aid.a <- aov(aid ~ GEN, data = aid)
summary(aid.a)
TukeyHSD(aid.a)


#Anova test between aid received
SNC$GEN <- LVI.m$GEN
aid.p <- aov(PCN ~ GEN, data = SNC)
summary(aid.p)
TukeyHSD(aid.p)
```
