

The Use of AI to Measure Gross National Happiness

Riona Krishnan Dighe

Cheltenham Ladies College, United Kingdom; riona.krishnan@gmail.com

ABSTRACT: This research attempts to identify an alternative approach to measuring Gross National Happiness (GNH). It uses artificial intelligence (AI), incorporating natural language processing (NLP) and sentiment analysis to measure GNH. It uses four widely used AI approaches to estimate GNH scores from online content such as social media posts and news websites to provide a cost-practical alternative approach to the traditional approach of GNH measurement. Traditional GNH measurement uses expensive and time-consuming surveys that take place infrequently, once every few years. This research addresses some of the limitations of conventional surveys. We use existing 'off the shelf' NLP models to encode social media-based text data and then conduct sentiment analysis, testing several AI classification approaches to estimate GNH scores. The four classification approaches tested to estimate the sentiment scores are the Multi-layer Perceptron (MLP) classifier, decision tree, linear model, and K- nearest neighbors (K-NN). The results show the MLP approach to achieve the highest accuracy on the test data set. The research concludes that this approach can estimate a country's GNH scores more cost-effectively and timely than the traditional approach. Applying AI on social media and news websites provides a real-time understanding of society's progress, concerns, and challenges. This can help cost-effectively influence effective policy-making. A pilot study was conducted to measure the GNH of the Royal Kingdom of Bhutan.

KEYWORDS: Artificial Intelligence, NLP, Sentiment Analysis, Gross National Happiness (GNH), A Pilot Study Measured the GNH of the Royal Kingdom of Bhutan in 2023.

■ Introduction

The traditional approach to measuring sustainable development and well-being of a population based on GNHI (Gross National Happiness Index) utilizes surveys, which are time-consuming and expensive and thus take place once every few years. Therefore, we conducted a desktop review to establish the use of AI to give real-time measurement of GNHI as a cost-effective alternative to the current approach.

GNHI is an emerging multi-factor index¹ that offers a holistic approach to measuring sustainable progress in society compared to the conventional measures using gross domestic product (GDP) and gross national product (GNP) data, which solely relies on financial and economic metrics.²

Some questions arise about the relevance of GDP/GNP measures given the accumulation of crises that have shocked the world: Financial, food, and energy crises contributing to increased cost of living and income disparities, including environmental disasters of the magnitude and frequency never seen before.^{3,4}

This is driving the adoption of more comprehensive and holistic methods to measure a country's sustainable progress and development using indices or attributes related to GNH.⁴ The Royal Kingdom of Bhutan was the first to define the concept in 1972, and GNHI was implemented as a tool for public policy in 2005.^{4,5} We further observe countries such as New Zealand and the United Arab Emirates adopting some attributes of GNHI in their policymaking. Further, META (Facebook) has used AI based on NLP to predict individual

well-being in social media.⁶ This has inspired me to conduct AI research to measure a country's GNH.

Research Approach:

Our research was conducted based on the nine target domains used to compute GNHI as defined by the Royal Kingdom of Bhutan: psychological well-being, good governance, education, health, ecological diversity, community diversity, cultural diversity and resilience, living standards, and time use.^{4,5}

A survey with questions specifically crafted based on the nine domains of GNH was undertaken. The nine questions are as follows:

Table 1: Sample survey crafted on the nine domains of GNH.

Domains	Questions
Psychological Well-being	Can you live a meaningful life in your current environment?
Good Governance	Are you happy with the process and institutions in place for policy decisions?
Education	Are you satisfied with the standard of education in the UK?
Health	Are you pleased with the standard of healthcare in the UK?
Ecological Diversity	Are you happy with the policies and direction of nature conservation efforts in the UK?
Community Vitality	Is the government successful in achieving prosperity and harmony amongst the different communities in the UK?
Cultural diversity and resilience	Are there adequate cultural awareness and education to sustain cultural diversity?
Living Standard	Are you content with your current living standard?
Time use	Do you think you are using your time efficiently when conducting activities with government agencies and the private sector?

The algorithm based on 'off the shelf' NLP models using sentiment analysis was tested against the answers from a small sample survey comprising thirty-five respondents in the UK. A simple score was derived from the algorithm to indicate the GNH of the sample. The scores generated by the algorithm closely resembled the human responses and were statistically

significant. This gave us the confidence to use NLP models against websites' sentences to measure a country's GNH. We conducted a pilot study using the NLP models to calculate the Royal Kingdom of Bhutan's GNH.

■ Pilot Study and Results

The NLP models constructed were applied to the English language social media of the Royal Kingdom of Bhutan, including newspapers. The attributes for each of the nine domains were reviewed, and the NLP models were applied. The characteristics of the nine domains are listed in Appendix 1. A raw sentiment analysis score (0 -10) was derived for each domain attribute, and a simple average was computed to produce the raw score for the domain. Finally, an overall raw score was obtained for the GNH of the Royal Kingdom of Bhutan using a simple average of the nine raw scores derived across the nine domains.

The following social media websites were used: kuenselonline.com, thebhutanese.bt, bhutantimes.bt, businessbhutan.bt, drukjournal.bt

The results are as follows:

Table 2: Raw sentiment analysis score derived from the NLP models applied to the English language social media of the Royal Kingdom of Bhutan. Results show high scores for cultural diversity, ecological diversity and psychological well-being.

Domains	Raw Sentiment Analysis Score
Psychological Well-being	8
Governance	7
Education	6
Health	7
Ecological Diversity	9
Community Vitality	7
Cultural Diversity	9
Living Standard	6
Time Use	8
GNH Raw Score	7

AI Algorithms:

We built an AI algorithm using NLP models that are responsible for giving a sentence a score between 0-10. A score of 0 corresponds to a very negative sentence, while a score of 10 indicates an extremely positive sentence. The algorithm was tested on responses by thirty-five randomly selected end-users in the UK. We observed that the scores generated by the algorithm closely resembled the human responses. The input to the algorithm is a sentence or a paragraph, and the output is a score between 0 and 10.

The algorithm consists of an encoder-decoder model, which has been used for sentiment analysis purposes. These models have been trained in understanding text and converting them into numbers, which are then given to a decoder, a neural network responsible for scoring the sentence. The neural network utilizes the vector indices as its input features, relying on predefined ground truth labels for its training process. This architecture and the pre-trained model were leveraged as a black box approach to generate sentiment analysis scores. Given the existing model's robustness and accuracy, no additional fine-tuning or further training is necessary.

■ Methodology Models

Before any AI model could be utilized, the data had to be pre-processed for the least amount of errors.⁷ Many individuals had no response to a specific question, which would skew the model to predict users' responses to a particular question inaccurately. Deleting these responses allowed for more efficiency in the data, leading to a more successful accuracy. Then, the data was split into train and test datasets. As mentioned before, 80% of the data was utilized for the training dataset, while the remaining 20% was used for the test dataset. Figure 1 demonstrates how data is split into test and train groups. A neural network was then fitted to the training set and utilized to predict the test question. This was done using a scikit-learn (sklearn) model library.⁸

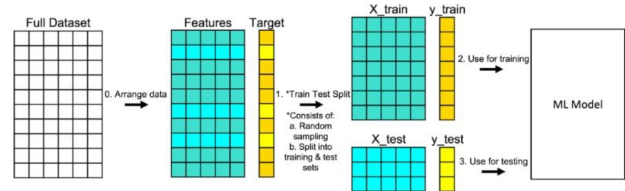


Figure 1: Demonstrate the splitting of data into test and train groups.⁹ 80% of the data is used for the training dataset and 20% for the test dataset.

The first classification model used was linear. Linear models are the simplest to use and try to generate a best-fit line to predict unknown values. In this case, a linear classification was used because I predict a numerical value - the answer a user may give from 1-5. Figure 2 represents how linear classification functions, dividing data points into classes. However, this model could perform better due to the characteristics of user responses. User responses to questions do not follow a trend and can be seemingly random, making choosing a linear model a poor choice; it obtained a test accuracy of 47.95%.

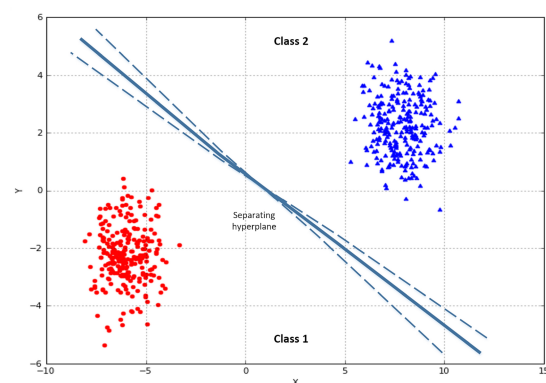


Figure 2: Example of how linear classification functions generate a best-fit line to predict unknown values.

The metric that was used to depict the results was an accuracy statistic. Given that the primary objective is to assess how well the model can predict the sentiment scores in general rather than focusing on specific class performance, accuracy provides a straightforward and comprehensive

measure of performance across all sentiment categories. Table 3 below summarizes the classification models used and their respective train and test accuracies. The Multi-layer Perceptron classifier ("MLP") had the best test accuracy, although the decision tree had the best train accuracy. Overfitting may have been an issue for this case's decision tree classification model, where its train accuracy is high but with low test accuracy, highlighting that the model has learned the patterns in the training data well but performs less well on new unseen data. However, this is not the case for the MLP approach, which has both high training and test accuracies.

Table 3: Various classifier models and results. MLP Classifier provides the best test accuracy.

Model	Train Accuracy (%)	Test Accuracy (%)
MLP Classifier	86.78	90.12
Decision Tree	96.85	54.75
K-nearest neighbors	58.75	52.50
Linear Classification	49.35	47.95

A hyperparameter was modified for the model using specific input questions. The choice of questions can be seen as a hyperparameter that can accurately predict a particular question. These questions were related to measuring GNH using the nine domains specified.

This model can perform poorly when the number of input questions is small. In addition, humans are unpredictable, and there is no de facto response to any question. People may answer questions differently at different points due to various circumstances and emotional states that can vary over time. For example, a gloomy person will make statements that appear damaging even when they speak about their most preferred outcome, and vice versa for a naturally optimistic person. All these can create outliers and affect the models' results. That said, the law of large numbers helps mitigate these effects. Using a large set of questions posed to a large enough sample size, these differences tend to average out, leading to more accurate results.

■ Conclusion

Our approach allows real-time analysis of GNH using AI to derive GNHI. Real-time GNHI can help cost-effectively influence effective policy-making and the government's economic, social, and environmental decisions. It can send early warning signals to policymakers and the government if key indicators begin to trend downwards and allow timely remedial action. Using AI on social media and news websites provides a real-time understanding of society's progress, concerns, and challenges.

However, some things could be improved in this AI-driven approach, which will affect the accuracy of the reported GNHI scores. A combination of factors, such as the existence of a free press, the extent to which the government censors the press and media, and levels of penetration of the internet and smartphones, can affect the results of the GNHI analysis. In addition, the results can be biased due to populist sentiment, including noise created by fake news. Further, the robustness of the scoring algorithm used in AI has a critical impact on the results of the reported GNHI:

Garbage In, Garbage Out. In addition, certain countries may have many languages, and the results would be unrepresentative of the entire population, as the algorithm used in the study is based only on the English language.

Further Enhancements:

The following steps of research can be extended to using a wider population for the surveys with the inclusion of more questions to test the accuracy of the scoring algorithm of the selected NLP models. The nine domains could be further investigated to determine whether they are the most representative of the measurement of the sustainable progress of a country. This may result in the addition or deletion of domains. An equal-weighted scoring methodology was used in the pilot study to determine the overall raw score for the GNH of the Royal Kingdom of Bhutan. A weighted average scoring method could be used based on the relative importance of the domains when determining the overall raw GNH score.

Further, the dataset could be extended to include a more comprehensive population dataset with AI-translated output covering many local and regional media and newspapers in the local languages. This will capture a broader population than the narrow subset that reads and writes English, thus creating social biases and distorting the results. ChatGPT or Copilot could be utilized against all social media sites, including English and Dzongkha in the Royal Kingdom of Bhutan, to derive answers to questions related to each attribute within the nine domains. The algorithm could then be applied to the answers from ChatGPT or Copilot to derive the raw sentiment score for each domain. This will capture the sentiment of the entire population across time. Bhutan is an important choice for this pilot study to be extended for further research analysis, given that the country has a database consisting of decades of manual surveys that can be tested against the results of the AI methodology.

■ Acknowledgments

I am grateful to my mentor, Odysseas Drosis, Ph.D., an alumnus of Cornell University, for his valuable guidance, suggestions, and criticism of the evolving research. Odysseas has inspired me in this journey towards understanding the various uses of AI. My reviewers provided valuable comments and suggestions, which have been incorporated into this research. I would like to extend my appreciation to the following reviewers:

Professor Prajit Dutta, Department of Economics,
Columbia University, United States of America

Dr. Dan Philips, Honorary Research Fellow, University of
Warwick. Head of Rothko Investment Strategies, London,
United Kingdom

Dr. Finian Tan, Founder & Chairman of Vickers Venture
Partners, Singapore.

Mr Manu Bhaskaran, CEO of Centennial Asia Advisors,
Singapore

■ References

1. Oxford Poverty and Human Development Initiative (OPHI), Bhutan's Gross National Happiness Index, University of Oxford.
2. Economics Third Edition, David Begg, Stanley Fischer, and Rudiger Dornbusch.

3. Beyond GDP Expert Recap – Genuine Progress, Aniket Shah & ESG Strategy Team - Jefferies (2022).
4. Gross National Happiness: Practice and Measurement, The Proceedings of the Fourth International Conference on GNH, 24-26 Nov 2008, The Centre for Bhutan Studies, Dasho Karma Ura and Dorji Penjore.
5. GNH WF Method and Optimization of GNH Resources, Mukund M Moharir (2016).
6. Predicting Individual Well-Being Through the Lens of Social Media, Proceedings of the Pacific Symposium, 2016, H. Andrew Schwartz, Margaret L Kern, Maarten Sap, Johannes C Eichstaedt.
7. The impact of preprocessing on text classification, Information Processing & Management, Volume 50, Issue 1, 2014, Pages 104-112, Alper Kursat Uysal and Serkan Gunal.
8. Classification with Learning K-Nearest neighbors, Proceedings of International Conference on Neural Networks, 1996, Pages 1480-1483, Volume 3, J.Laaksonen and E.Oja.
9. Understanding Train Test Split, Built In, <https://builtin.com/data-science/train-test-split>.
10. Commission on Measurement of Economics Performance and Social Progress, JE Stiglitz, A Sen and J-PFitoussi (2008).
11. A Well-Being Manifesto for a Flourishing Society, Nic Marks and Hetan Shah (2004).

■ Author

Riona Krishnan Dighe is a Year 12 (UK) full-time boarding student at Cheltenham Ladies College, United Kingdom. She has been inspired by the growing use of AI across various industries, including policy-making. She personally presented her research and pilot study results to the Centre for Bhutan and GNH Studies (CBS) in Thimphu, Royal Kingdom of Bhutan.

■ Appendix 1

There are nine domains in the GNH.

We use the AI algorithm to find the sentiment related to several indicators concerning each domain.

A) Education

- 1) Cultural Literacy
 - Frequency of Reporting and discussions on cultural and religious festivals
 - Participation by schools in local cultural festivals
- 2) Political Awareness
 - Frequency of free and open discussions on politics, constitutional rights, national council, national assembly
 - Discussions around Gross National Happiness
- 3) Environment
 - Frequency of Discussions related to climate change, mitigation and solutions
 - Words related to hydropower, solar power, wind farm
- 4) Healthy Living
 - Awareness and discussions around healthy living, diet, exercises, well-being and mental health
- 5) Skills-based education
 - Availability and popularity of vocational education: weaving, carpentry, embroidery, painting, carving, sculpture, casting, blacksmithing, bamboo works, masonry, leather work, paper making, electrical maintenance, automobile technicians
- 6) Formal education
 - Literacy level - reporting and discussions
 - Comparison between urban and rural/village
 - Comparison between male and female
- 7) Satisfaction with:
 - Content of courses
 - Preparing for employment

- Quality of teachers
- Student-teacher ratio
- Quality of educational infrastructure:
- Classroom size/lighting/ventilation/furniture
- Toilet adequacy and cleanliness
- Library adequacy
- Science laboratories
- Computer laboratories
- Sports facilities
- Safe drinking water
- Sufficient and good condition of textbooks
- Availability of special needs education: blind, deaf, and disabled children
- 8) Reporting on Vices
 - Smoking
 - Alcoholism
 - Drug Abuse
 - Underage marriage and pregnancy
- 9) Unemployment
 - Unemployment rate and discussions
- 10) Gender Inequality
 - Discussions around gender inequality

B) Health

- 1) Discussions on general health, mental health, disability problems, and activity limitations
- 2) Reporting on:
 - Alcoholism
 - Smoking
 - Chewing tobacco
 - Chewing khamto
 - Drug abuse
- 3) Reporting of suicides and suicide rate
- 4) Satisfaction with:
 - Quality of healthcare
 - Hospital services
 - Skill and competency of staff
 - Cleanliness of hospitals
 - Friendliness and courtesy of the staff
 - Treatment and medicine prescribed
- 5) Confidence and satisfaction with healthcare access
 - Discussions on the availability of healthcare services
 - Discussions on distance to healthcare facilities
 - Discussions on the availability of transportation to healthcare facilities
 - Discussions on waiting times to see doctors

C) Psychological Well-Being

- 1) Satisfaction with the following aspects of life:
 - Health
 - Standard of living
 - Occupation
 - Relationship with family
 - Local environment
 - Work-life balance
 - Meaning in life
 - Freedom of expression
 - Feeling accomplished
 - Feel cared for in life
- 2) Social Support Provision:
 - When sick
 - In financial difficulties
 - When suffering from emotional problems
 - When need to attend important personal events (childbirth, funeral, wedding)
- 3) Stress Levels

- Frequency of Reporting and discussions around stress and strain
- 4) Happiness
 - Discussions around happiness and a sense of worth
 - 5) Unhappiness
 - Discussions around unhappiness and depression
 - 6) Quality of Life
 - Discussions around the enjoyment of life
 - 7) Spirituality
 - Words related to spirituality, prayers, meditation, spiritual discussions, local temple activities, Karma
 - 8) Emotional experience
 - Words related to anger, guilt, selfishness, jealousy, pride, compassion, forgiveness, calmness, contentment, generosity, sadness, frustration, fear, worry
- D) Cultural Diversity and Resilience
- 1) Frequency of reporting and discussions on cultural traditions, local sports, and religious festivals
 - 2) Discussions around the use of mother tongue
 - 3) Importance of:
 - Family life
 - Friendship
 - Generosity
 - Spiritual faith
 - Compassion
 - Self-development
 - Reciprocity
 - Responsibility
 - Freedom
 - Material wealth
 - Career success
 - Material wealth
 - Pleasure
 - 4) Public opinion on:
 - Killing
 - Stealing
 - Lying
 - Harsh and offensive speech
 - Harmful thoughts
 - Sexual misconduct
 - False views
 - 5) Discussions on educating children towards:
 - Respect for elders
 - Respect for parents
 - Discipline
 - Honesty
 - Tolerance
 - Hard work
 - Obedience to authority
 - Caring for family members
 - Helping others
 - 6) Awareness of gender equality
 - Discussions related to the importance of education for boys instead of girls
 - Discussions related to women being more suited for domestic chores than men
 - 7) Popularity of astrologers
 - 8) The popularity of Hindi and English movies compared to Bhutan movies
 - 9) The popularity of traditional architecture for houses and buildings
- E) Governance
- 1) Satisfaction with the Performance of local and national government concerning:
 - Creating jobs
 - Reducing income inequality
 - Provision of educational needs
 - Improving health services
 - Fighting corruption
 - Protecting environment
 - Improving infrastructure
 - Providing farm and feeder roads
 - Improving sanitation
 - Providing potable water (clean drinking water)
 - Providing stable electricity
 - Providing agricultural implements and seeds
 - Approving allocation of timber permits
 - Preserving culture and traditions
 - 2) Perception of independence of the judicial system
 - Fairness, speed, and affordability of the legal system
 - 3) Fair electoral process
 - 4) Fundamental Rights
 - Access to freedom of speech and opinion
 - Right to vote
 - Right to join a political party of choice
 - Right to equal pay for work of equal value free from race, gender, religion, language, or political discrimination
 - 5) Perception of law and order by the police force
 - Satisfaction of law enforcement
 - Satisfaction of crime prevention
 - Satisfaction with police response time
 - The police force is approachable and easy to talk to
 - 6) Free and open press (radio, TV and newspapers)
 - 7) Transparency, accountability, and honesty in government institutions e.g., courts, police, ministries, local government, national assembly national council
 - 8) Corruption
 - Reporting and discussions on:
 - Misuse of government money
 - Misuse of subordinate staff
 - Misuse of government vehicles and equipment
 - Favoritism in employment and promotion
 - Favoritism in the award of contracts
 - Favoritism in issuance of trade licenses for mining, timber, sand, or quarry
 - Corruption in the procurement of goods and services
 - Corruption in budget allocation
 - Buying Votes using money or gifts
- F) Community Vitality
- 1) Discussions on volunteering:
 - Voluntary labor contribution to religious rituals
 - Voluntary labor for house construction
 - Voluntary labor for construction/repair of irrigation channel
 - Voluntary labor for house repair
 - Voluntary labor contribution during times of death in a family
 - Voluntary labor contribution during paddy plantation and harvest
 - Voluntary clean-up campaign
 - Voluntary teaching
 - Voluntary fundraising
 - 2) Discussions around donations:
 - To religious establishments
 - To religious rituals
 - To other families
 - To educational institutions
 - To NGOs

-
- 3) Sense of Belonging and Trust
 - Level of trust and sense of belonging in the local community
 - 4) Community Spirit
 - Strong community interaction and participation
 - 5) Strong Family Bonds
 - 6) Crime and Safety
 - Discussions around theft, robbery, family violence, fraud, vandalism and sexual assault
 - Feeling safe walking alone in the neighborhood
 - G) Ecological Diversity
 - 1) Environmental Knowledge and Awareness
 - Discussions related to:
 - Pollution
 - Anti-pollution laws
 - Environmental standards
 - Taxes related to companies importing environmentally damaging goods
 - 2) Pollution
 - Discussions around:
 - Pollution of rivers and streams
 - Air pollution
 - Noise pollution
 - Littering
 - 3) Recycling and Waste management
 - Discussions around the absence of waste management
 - Discussions on reusing and recycling materials, e.g. plastic bottles, plastic bags, tins
 - Discussions around
 - Reduction of waste
 - Segregation of waste
 - Sorting biodegradable and non-biodegradable waste and disposal in special containers
 - Discussions around the disposal of household waste such as burning, dumping in rivers, dumping in forests, dumping on open land
 - Positive discussions around household waste using composting and municipal pick up
 - 4) Energy
 - Buying energy-efficient bulbs and appliances, keeping the TV, computer, or lights on
 - Renewable energy sources, e.g. hydroelectricity, solar power, and wind farms
 - Heating homes using electric heaters compared to polluting kerosene heaters or wood-burning
 - 5) Conservation of water
 - Not leaving the water tap on
 - 6) Environmental Disasters
 - Frequency of discussions around landslides, soil erosion, floods, erratic weather patterns
 - 7) Deforestation
 - Decreasing plant species
 - Decreasing wild animals
 - 8) Urban Environmental Issues
 - Decreasing air quality
 - Traffic congestion
 - Absence of green spaces
 - Lack of pedestrian-friendly streets
 - Unplanned urban sprawl
 - 9) Eco-friendly behavior
 - Travel sustainably by walking, cycling, or using public transport
 - Participate in clean-up activities
 - 10) Agricultural land lost to development
 - 11) Water supply
 - Discussions around access to water supply
 - Quality of drinking water
 - Discussions around discoloration, odor, sediment, and taste of drinking water
 - 12) Satisfaction with public transport
 - Availability, Accessibility, Affordability, Cleanliness, Timeliness and frequency of public transport
 - 13) Climate change
 - Awareness of climate change and mitigation
 - 14) Agriculture
 - Quality of agricultural land
 - Discussions around organic farming
 - Use of organic fertilizers, natural pesticides such as vinegar
 - Negative discussions around the use of chemical fertilizers or inorganic fertilizers and inorganic pesticides
 - 15) Agriculture Constraints
 - Discussions around crops destroyed by:
 - Wild animals
 - Domestic animals
 - Pest and diseases
 - Drought
 - Excessive rain
 - Hailstorm
 - Soil erosion
 - Landslides
 - Labor shortage
 - Lack of stable water supply/irrigation system
 - Late supply of approved seed by agency
 - Lack of fertilizer or manure
 - H) Living Standard
 - 1) Income and Affordability
 - Discussions around income sources
 - Discussions around income inequality
 - Discussions around change in financial position
 - Discussions around financial security
 - Discussions around the affordability of necessities. e.g. food, clothing, and shelter
 - Discussions around the lack of savings available for emergency expenditure
 - Discussions around the lack of affordability for children's education
 - Discussions around the lack of ability to repay loans and mortgages
 - Discussions around difficulty in providing financial assistance to parents and family members
 - 2) Discussions around household and business indebtedness
 - 3) Homeownership levels in Bhutan
 - 4) Average number of rooms in homes
 - 5) Availability of toilets and flushing toilet facilities in homes
 - 6) Availability of stable electricity in households
 - 7) Availability of stable potable water supply to homes
 - 8) Percentage of homes using bricks and concrete compared to wood, mud, and straw leaves
 - 9) Asset ownership in Bhutan e.g. land ownership
 - 10) Livestock ownership in Bhutan
-