Demographic Research

Training Manual

Introduction

This Demographic Research Training Manual equips mission strategists, national reps, and researchers with practical methods to estimate the Christian and evangelical population within unreached people groups (UPGs). Given the challenges of conducting full population censuses, this guide focuses on three key methods:

1. **Network-Based Estimation:** Using denominational and church data to estimate Christian populations.
2. **Survey Sampling:** Collecting statistically valid data across representative locations.
3. **Informed Observation:** Gathering expert consensus when direct data is limited or unavailable.

Designed for individuals with limited statistical training but strong cultural understanding, this manual provides step-by-step instructions for data collection, verification, and bias reduction. It emphasizes combining multiple methods to achieve reliable estimates and offers practical tools, including sample size guides, survey templates, and verification techniques.

By following this guide, you can generate accurate data to inform mission strategies, adapt research to local contexts, and refine estimates over time.

Note about Christian Adherent and Evangelical designation

In this manual the term Christian and Evangelical is used interchangeably, but they do represent **two distinct categories** that Joshua Project seeks to quantify and present. In highly unreached contexts these two numbers are typically very close.

Christian Adherent: Anyone who professes to be Christian. The term embraces all traditions and confessions of Christianity. It is not an indicator of the degree of commitment or theological orthodoxy. This definition is based on the individual's self-confession, not his or her ecclesiology, theology, denomination or religious commitment and experience.
Because Joshua Project works toward total population in a country, Christian adherent estimates include entire families including all children.

In contrast, **Evangelicals**, are followers of Christ who generally emphasize:

- The Lord Jesus Christ as the sole source of salvation through faith in Him. Personal faith and conversion with regeneration by the Holy Spirit. A recognition of the inspired Word of God as the only basis for faith and living. Commitment to Biblical preaching and evangelism that brings others to faith in Christ.

The noun "Evangelical" is capitalized since it represents a body of Christians with a fairly clearly defined theology (as also Orthodox and Catholic bodies, etc.). Evangelicals are here defined as:

- All affiliated Christians (church members, their children, etc.) of denominations that are evangelical in theology as defined above.
- The proportion of Christians in other denominations (that are not wholly evangelical in theology) who would hold evangelical views.
- The proportion of Christians in denominations in non-Western nations (where doctrinal positions are less well defined) that would be regarded as Evangelicals by those in the above categories.
- This is a theological and not an experiential definition. It does not mean that all Evangelicals as defined above are actually born-again. In many nations only 10-40% of Evangelicals so defined may have had a valid conversion and also regularly attend church services. However, it does show how many people align themselves with churches where the gospel is being proclaimed.

Method 1: Denominational and Network-Based Statistics

Overview

This method involves collecting data from churches, denominational networks, and mission organizations that track church membership, baptisms, professions of faith or congregation sizes. By extrapolating available figures against the total population of the people group, an estimate of Christian percentage can be determined. This approach is particularly useful in contexts where a strong church presence already exists, and historical data can provide insight into growth trends.

1. **Gather Church Data**
    a. Request membership data from denominational leaders, mission agencies, or church networks.
    b. data examples:
        i. Baptisms over the past (available number of) years.
        ii. Total active church membership.
        iii. Generation maps of churches.
        iv. Average weekly attendance.
        v. Number of congregations.
        vi. Average size of congregations.
        vii. Growth trends based on historical data.
    c. **Note**: For data coming from Movement networks there is a special workflow that will be implemented for data incorporation.
2. **Aggregate Data**
    a. Baptisms/membership/attendance/etc.
        i. Combine data from various networks, accounting for possible overlap.
    b. Church Count
        i. Determine the mean or median congregation sizes.
            1. The mean is better when the data is symmetrical, while the median is better when the data is skewed or has outliers.
        ii. Example Calculation:
            1. If 140 churches are counted and a mean (average) membership of 11, the total estimated membership = 140 × 11 = 1,540.

3. **Accounting for Non-Affiliated Christians, Non-Reporting Churches, and Duplicated Reporting (Optional)**

In some cases, church records do not cover the entire Christian population within a people group. Some churches may not report membership data, and some believers may not be formally affiliated with a congregation. Similarly, some believers are churches may be counted multiple times from different groups. To generate a more accurate estimate we apply a **correction factor** to adjust for missing data. Always verify with **multiple sources** before applying adjustments.

a. **Adjust for Non-Reporting Churches**

If denominational data only accounts for a portion of known churches, an **extrapolation factor** is applied to estimate the total Christian population. The formula used is:

Estimated Total Christians = Reported membership/coverage percentage

For example:

1. If church data covers **70%** of known churches, apply a correction factor of **1.43** (1 / 0.7) to estimate the total population.
2. If **reported membership** is **1,540**, the adjusted estimate is:
    a. 1,540 x 1.43 = 2,202

This means the estimated total Christian population, accounting for non-reporting churches, is **2,202**.

b. **Account for Non-Affiliated Christians**

Not all believers are part of a formal congregation. Some may practice their faith privately due to personal choice, persecution, or lack of a local church. Based on **key informant interviews**, an adjustment can be made to include this population.

1. If local leaders estimate that **10%** of Christians are unaffiliated, apply a multiplier of **1.1**:
    a. 2,202 × 1.1 = 2,422

Thus, after adjusting for both **non-reporting churches** and **unaffiliated believers**, the total estimated Christian population is **2,422**.

c. **Account for Duplicated Reporting**

If duplication is suspected (e.g., same individuals or churches are counted by multiple sources), adjust by applying a conservative correction factor.

1. Aggregated numbers of Christians x (percentage of estimated duplication) = number of Christians

For Example:

1. If there is a likelihood of 50% duplication of data, apply a correction factor of **0.5** to estimate the total population.
2. If **reported number** is 2,422, the adjusted estimate is:
    a. 2,422 x 0.5 = 1,211

4. **Calculate estimated Christian Percentage**

Once you have adjusted for potential under or overcounts, calculate the percentage of Christian/Evangelical.

(Total Christians/ Population of the People Group in country) ×100 = Christian Percentage

a. Example:
    i. (2,422/600,000) x 100 = 0.4%

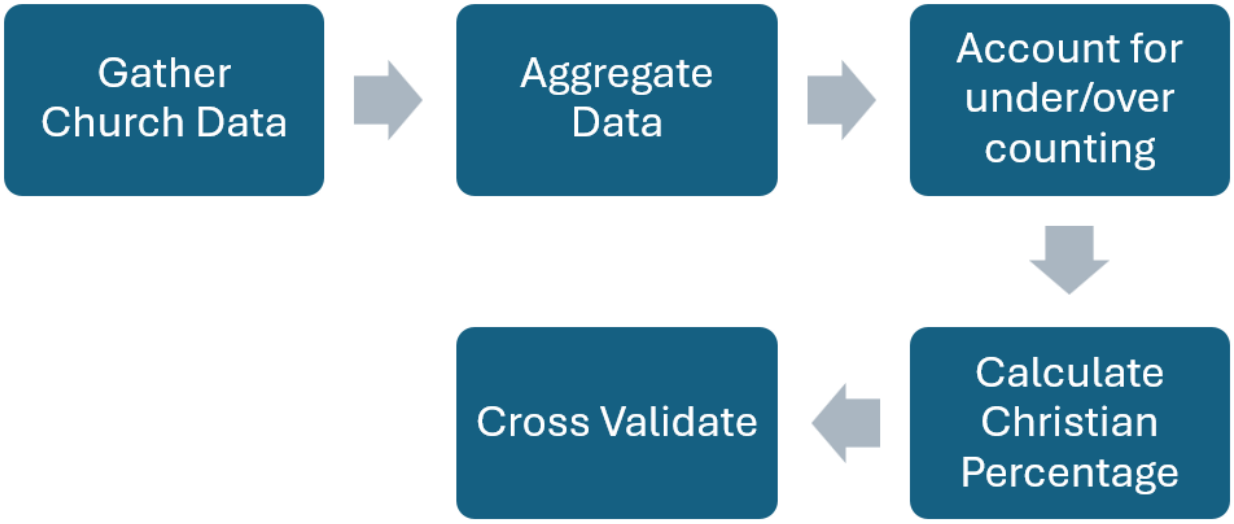
5. **Cross-Validate with Historical Trends**
    a. Compare with previous estimates of growth rates in the church.
    b. If there is significant deviation from historical trends, investigate again potential undercounting or overcounting biases.
    c. Be aware that church attendance rates vary widely across regions and cultural contexts.
    d. Some denominations may have a higher proportion of active members, while others have nominally affiliated members who do not regularly participate.

Strengths and Limitations

- **Strengths**:
    - Uses existing, verifiable data.
    - Less resource-intensive than surveys.

    - Avoids drawing attention to research activities in sensitive areas.
- **Limitations**:
    - Relies on accurate and complete church records.
    - May undercount non-affiliated Christians.
    - Cannot capture new converts or secret believers accurately.

Visual Flowchart



Method 2: Survey Sampling for Statistical Estimation

Overview

This method involves conducting surveys among a **representative sample** of a people group to estimate the Christian population with at least **90% confidence and a ±1% margin of error and ideally 95% confidence and a ±0.5% margin of error.** This method is helpful where there is an emerging church or a church that has not institutionalized much for widespread data collection.

1. **Define the Target Population**
    a. Identify the total population of the people group.
    b. Ensure urban and rural areas are represented.

2. **Determine the Sample Size**

Link to Sample Size Calculator

If viewing printed version or if web access isn't available, see appendix B to calculate minimum sample size manually, or estimate based on the table in appendix B.

3. **Stratify Survey Locations**
    a. **Stratified sampling** is a method where the population is divided into **distinct subgroups or strata** based on characteristics such as **socioeconomic status, age, gender, education level, or geographic location**. The goal is to ensure that each subgroup is properly represented in the sample according to its proportion in the total population.

How Stratified Sampling Works:

1. **Divide the Population into Strata:**
    a. For example, if a population includes both urban and rural residents with varying income levels, you might create strata like:
        i. Urban, high-income
        ii. Urban, low-income
        iii. Rural, high-income
        iv. Rural, low-income

2. **Determine Proportional Representation:**
    a. Calculate how much each stratum represents in the total population. For example, if 30% of the population falls into the “Urban, low-income” stratum, then 30% of the survey sample should be drawn from this group.
3. **Sample Within Each Stratum:**
    a. Randomly select participants within each stratum to avoid bias, ensuring that each group contributes proportionally to the overall sample.

4. **Conduct the Survey**
    a. Ask neutral, clear questions (e.g., "What is your religious affiliation?").
    b. Use local language and cultural sensitivity.
    c. Ensure responses remain anonymous.
    d. Conduct pilot testing to refine survey questions and methodology.
    e. See **Appendix C**: for a S**ample Community Survey** for examples in building your survey.

5. **Calculate the Percentage of Christians**
    a. Count the number of Christian responses compared to total number surveyed.
    b. Use the formula:
        i. Christian Percentage = (Christian Responses/Total Surveyed) x 100
    c. Link to **Christian Percentage Calculator**

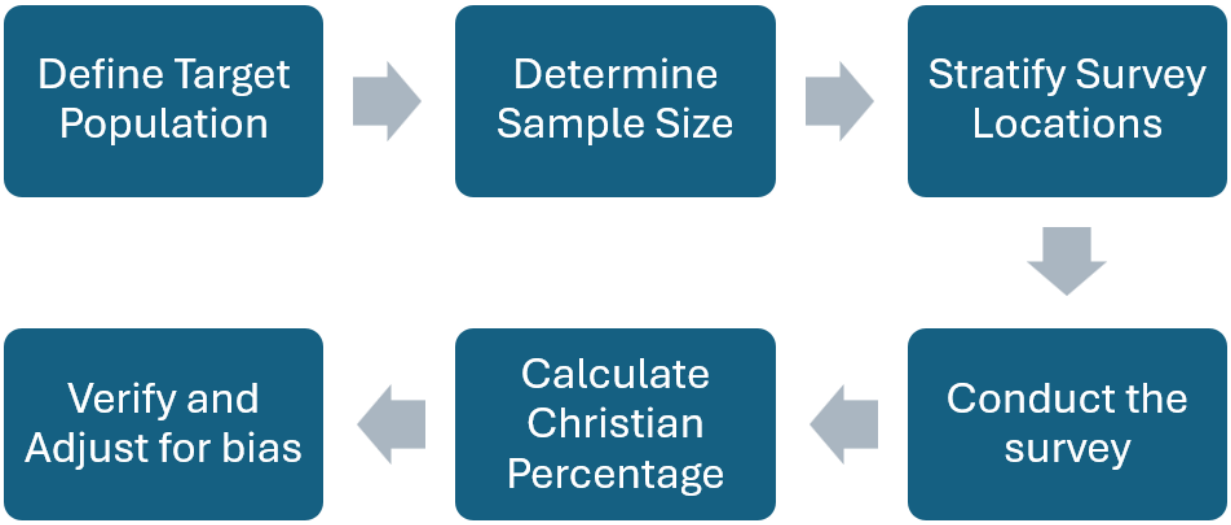
6. **Verify and Adjust for Bias**
    a. Compare results with known denominational, country, network, etc. Data.
    b. Consider factors like underreporting due to persecution or cultural reluctance to disclose religious identity.
    c. Cross-check against historical trends or census reports.

7. **Calculate range based on confidence interval (Optional)**
    a. Link to **Range of Christian percentage Calculator**

Strengths and Limitations

- **Strengths**:
  - High accuracy when sample size requirements are met
  - useful in regions without robust denominational data.
- **Limitations:**
  - Requires significant resources
  - possible underreporting due to sensitivity or reluctance.

Visual Flowchart



Method 3: Informed Observation: Gathering Consensus Through Local Experts

There are times that there are several well informed “experts” who are knowledgeable about the status of Christian witness among a people group. These people may be cross-cultural workers, indigenous leaders, or regional researchers. These experts can be questioned and consensus built through the Delphi method.

Overview of the Delphi Method

The Delphi method is a structured communication technique originally developed to reach consensus through a series of questionnaires or interviews with local experts. It is ideal for scenarios where direct survey data may be unavailable or where expert consensus is necessary to validate other estimates. This method allows for cross-checking data collected from surveys and denominational reports, creating a well-rounded understanding of Christian presence within a people group.

When to Use the Delphi Method

- When survey data is difficult to obtain or unreliable.
- When significant regional variations exist within a people group.
- When cross-verifying estimates gathered through surveys or denominational reports.

Steps for Conducting a Delphi Survey

1. **Identify Local Experts.**
    a. Seek out church leaders, mission workers, denominational representatives, indigenous leaders, and other informed individuals who have **experience with the target people group.**
    b. Ensure a diverse range of experts representing **different regions** within the country.

2. **Develop Initial Questions:**
    a. Begin with open-ended questions to gather baseline estimates and general perceptions:
        i. "What percentage of the people group do you estimate to be Christian?"

ii. Out of 1,000 people how many are followers of Jesus?
iii. "What percentage would you classify as evangelical?"
iv. "Are there regions within the group where Christian presence is higher or lower?"

3. **Round One: Initial Survey/Interview**
    a. Conduct the first round of questions, collecting responses from all identified experts.
    b. Emphasize that responses should reflect their best knowledge or observations and not guesses.

4. **Analyze and Summarize Responses:**
    a. Identify areas of agreement and divergence.
    b. Provide a summary of responses to all participants, highlighting key areas where consensus is strong or weak.

5. **Round Two: Feedback and Revision**
    a. Share the initial findings and allow participants to revise their estimates based on feedback and additional information from other experts.
    b. Encourage discussions on why differences exist and what factors could influence their estimates.

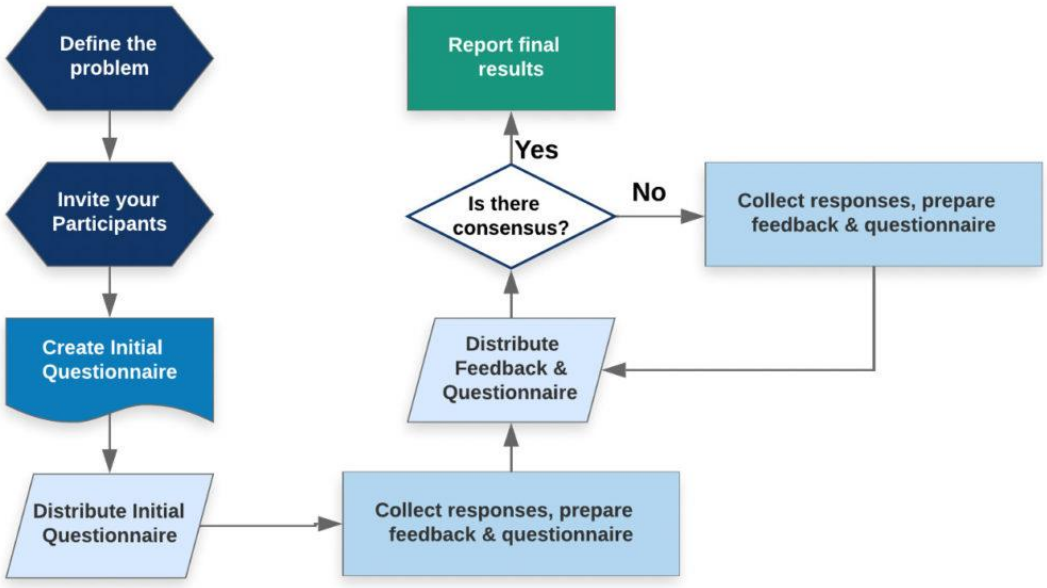
6. **Final Round: Consensus-Building**
    a. Conduct a final round of surveys, aiming for a consensus or reasonable range of estimates.
    b. If consensus cannot be reached, document the range of estimates and reasons for divergence.

7. **Final Report:**
    a. Compile the findings into a final report, documenting key insights, consensus ranges, and regional variations.

Strengths and Limitations

- **Strengths**:
  - Useful when direct data is scarce
  - allows cross-verification of other methods.
- **Limitations:**
  - Subject to expert bias
  - may lack statistical precision if consensus is weak.

Visual Flowchart



Collecting and Verifying Data

Accurate data collection and verification are essential for estimating the Christian and evangelical population within a people group. This section outlines practical tips for collecting reliable data, cross-verifying information from different sources, and mitigating bias.

Practical Tips for Collecting Christian/Evangelical Percentages

Use Neutral, Clear Wording:

- Frame questions about religious affiliation in a way that is unbiased and non-accusatory to avoid influencing responses.
    - *Example:* Instead of "Are you a Christian?" ask, "What is your religious affiliation?"

Ask for Ranges:

- When informants are unsure or reluctant to give precise percentages, ask them to provide a range.

Verify Ethnicity:

- In areas with diverse populations, it is crucial to ensure informants **correctly identify as the people group being studied**. This is particularly important in regions where a shared trade language may obscure ethnic distinctions. Structure your survey questions to differentiate between ethnic and linguistic identities.
    - *Example:* Instead of only asking about religious affiliation, include follow-up questions to clarify ethnic identity when needed.

Convert to Concrete numbers:

- If large discrepancies exist between local experts convert percentages into concrete numbers for better clarity.
    - **Example:** If the population of a group is 38 million and a leader reports 10% Christians (3.8 million), ask, "Are there more or less than 3 million believers?" This helps clarify estimates when informants realize the scale of the reported figure.
    - **Example:** Are there more or less than 100 believers out of a village of 1,000?

Consider Regional and Cultural Variations:

- Recognize that informants' knowledge may be localized. When conducting interviews, adjust your questions to capture any regional or tribal differences in Christian presence or religious dynamics.
    - The language used to identify religions may vary by context. Some may not identify as Christian, but as a term such as follower of Jesus.

Pilot Test Your Questions:

- Conduct initial interviews with a small sample group to refine question phrasing and survey structure. This helps identify and correct any potential issues before large-scale data collection.

Compare with Local Observations

- Check for differences between reported estimates and observable indicators, such as church gatherings, visible Christian symbols, or other signs of religious practice.
- Use informal interactions to uncover hidden or secret believers who may not be captured through formal surveys.

Trust-Building Techniques for Reliable Data Collection

Building trust with informants is key to obtaining honest responses, especially when discussing sensitive topics like religious identity.

- **Start with General Questions:** Begin interviews with non-threatening questions about community life or history to build rapport before discussing religious matters.
- **Allow Casual Conversations:** Let informal discussions flow naturally, which can make informants more comfortable.
- **Utilize Cultural Helpers:** Involve local individuals who are trusted by the community to assist in interviews and data collection.
- **Avoid Financial Incentives:** Offering financial rewards can lead to biased responses. Reassure informants that the data is for understanding the community and guiding mission efforts.
- **Future Considerations:** Let mission and church leaders know that their responses will help guide efforts to bring beneficial resources in the future. (e.g., outside workers, prayer networks).

Ethical and Security Principles

- **Anonymity and Confidentiality:**
    - Maintain the anonymity of participants, especially in regions where identifying as Christian may lead to persecution.
    - In documentation, do not record information that can identify people by those who would seek to harm.
- **Voluntary Participation:**
    - Participation should always be optional. Do not pressure or coerce individuals into responding.
- **Transparency:**
    - Have a short truthful statement prepared to define what you are doing. The statement should be truthful, limited in scope to protect you and your partners, and be easily remembered.
        - Example: *"We are gathering general community information to understand local demographics and trends."*

Submitting Updated Information

If submitting an update, include details such as the source of data, collection method, and any relevant adjustments made.

- **Joshua Project Web Submissions** – Each Joshua Project profile page has a **"Submit Update"** button. This is best done for individual people groups. You can submit updated information directly through that button for:
    - Data/Statistical updates
    - People Group Photos
    - People Group Maps
    - People Group Profile Text
- **Online People Group Research Form** - This form is great to share with others and has questions for untrained researchers to provide sufficient information for updating people group data. (Link to Form)

(Joshua Project National and Regional Reps Only)

- **Updating statistical information by spreadsheet** – In Microsoft Teams, your supervisor will create an Excel spreadsheet with each people group in your assigned country(s). There are added columns highlighted in yellow for you to provide statistical updates.
- **Uploading notes and files directly to Teams –** In Microsoft Teams, you will have a personal Uploads folder where you can upload field notes, pictures, resources, etc. After uploading, communicate with your supervisor what you uploaded and what action needs to be taken with the information.

By following these practices, you ensure that the knowledge you gather remains accessible and organized, helping you draw meaningful conclusions and identify areas for further exploration.

Conclusion

This manual provides three robust methods for estimating the Christian and evangelical population within an UPG. By leveraging existing denominational data and well-structured survey sampling, national reps can generate reliable and credible estimates to aid mission strategy and planning.

Additionally, we encourage continuous improvement in data collection, verification, and adjustment to refine future research efforts.

For further assistance, contact our research team for guidance and verification of collected data. **Happy researching!**

Appendix A: Glossary of Key Terms

1. **Bias:** A preference or prejudice that can affect the accuracy of data collection or results. In research, bias can come from the way questions are asked or how participants respond.
2. **Confidence Interval (CI):** A range that shows how certain we are about a result. For example, a 95% confidence interval means that we are 95% sure the true value lies within the given range.
3. **Cross-Verification:** The process of checking data or information from one source by comparing it with data from other reliable sources.
4. **Delphi Method:** A research method where experts provide estimates through multiple rounds of surveys, helping to build consensus on uncertain topics.
5. **Denominational Data:** Information gathered from churches or religious groups, such as the number of members, baptisms, or congregations.
6. **Extrapolation:** Estimating unknown values by using known data. For example, if data is available for 70% of churches, an estimate can be made for the entire group.
7. **Informed Observation:** Collecting information by observing a situation or by consulting people with knowledge of the subject, often without formal surveys.
8. **Margin of Error:** The amount of error allowed when estimating data. For example, a 1% margin of error means the estimate could be off by plus or minus 1%.
9. **Population:** The total number of people within a group or area being studied. For example, the population of a UPG includes all members of that group in a specific region or country.
10. **Sample Size:** The number of people surveyed to represent the entire population. Larger sample sizes usually give more accurate results.
11. **Statistical Estimation:** Using data from samples to estimate characteristics of the entire population. For example, estimating the percentage of Christians in a region based on surveys.
12. **Stratified Sampling:** A sampling method where the population is divided into subgroups (such as age or location) to ensure all groups are properly represented in the survey.
13. **Survey Sampling:** A method of gathering information by asking questions to a group of people selected from the population.
14. **Unreached People Group (UPG):** A group of people with little or no access to the Christian message due to cultural, linguistic, or geographic barriers.

Appendix B: Calculating and Estimating Sample Size

[Link to Sample Size Calculator]

When calculator is not available, you can use one of the below methods.

To achieve **xx% confidence with a ±x% margin of error**, use the formula:

Where: $$n = \frac{(Z^2) \times p \times (1-p)}{E^2}$$

- **n** = Required sample size
- **Z** = Z-score corresponding to the confidence level
  - for 95% confidence, **Z = 1.96** (preferred)
  - for 97% confidence, **Z= 2.17**
  - for 99% confidence, **Z = 2.57**
- **p** = Current estimated Christian percentage in decimal form
  - use **0.5** (50%) if unknown or previous data is unreliable
- **E** = Margin of error
  - **0.01** for 1%
  - **0.005** for 0.5%
  - **0.001** for 0.1%

To ensure reliable data, aim for a 95% confidence level with a ±1% margin of error.

How to use the table:

- Identify the sample that is closest to your situation. (population and current estimated percent Christian)
- Determine your desired Confidence Interval and margin of error
- Find the minimum sample size needed in the right column.
    - This table is only an example and should be used only when access to a sample size calculator isn't available.

Quick Reference Table for Sample Size examples:

Population of People Group	Estimated Percent Christian (%)	Z-Score	Margin of Error (Decimal)	Minimum Sample Size
10,000	1	1.96	0.01	367
100,000	1	1.96	0.01	379
1,000,000	1	1.96	0.01	381
10,000,000	1	1.96	0.01	381
10,000	5	1.96	0.01	1544
100,000	5	1.96	0.01	1793
1,000,000	5	1.96	0.01	1822
10,000,000	5	1.96	0.01	1825
10,000	0.1	1.96	0.005	152
100,000	0.1	1.96	0.005	154
1,000,000	0.1	1.96	0.005	154
10,000,000	0.1	1.96	0.005	154
10,000	0.5	2.17	0.005	857
100,000	0.5	2.17	0.005	929
1,000,000	0.5	2.17	0.005	937
10,000,000	0.5	2.17	0.005	937
10,000	10	2.17	0.01	2977
100,000	10	2.17	0.01	4066
1,000,000	10	2.17	0.01	4221
10,000,000	10	2.17	0.01	4237

Note about Minimum Sample Size:

- When the **proportion of Christians (p)** is very low (e.g., 0.1%), the variability in responses is relatively small. As a result, the number of required respondents to achieve meaningful accuracy decreases compared to cases where the population has a higher or unknown proportion of Christians.
- The **total population size has** minimal influence once it exceeds a few thousand people. For populations of this size and larger, sample sizes remain consistent for given confidence levels and margins of error. This is why the sample size is relatively small and does not scale proportionally with the population.
- From a practical perspective, **surveying a small subset** reduces the time, cost, and complexity of the research while still maintaining a statistically acceptable level of precision. The law of large numbers indicates that the sample average will approximate the true population proportion if the sample is random and independent.

Appendix C: Sample Community Survey

Section 1: Geographic Information

1. **State of Residence:**
2. **Local Government Area (if applicable):**
3. **Community or village Name:**

Section 2: Respondent Information

4. **Estimated Age:**

[Under 18, 18-24, 25-34, 35-44, 45-54, 55-64, 65+]

5. **Gender:**

[Male, Female]

Section 3: Ethnic/language Identity

6. **What is your primary ethnic identity?**

7. **Do you consider yourself a part of the primary ethnic group in this region?**

8. **What language do you speak at home?**

9. **Is that the language your parents speak?**

Section 4: Community Composition

10. **In your community, out of 1,000 people how many are of your ethnic group?** (ask only if the area is heterogeneous)
11. **In your community, out of 1,000 people how many speak your mother tongue language?**

Section 5: Religion

12. **What is your religion?**

a. Use culturally appropriate terms (i.e. follower of Jesus vs. Christian)

13. **In your community, how many (primary religion houses of worship) mosques/temples/wats are there?**

14. **In your community, are there other places of worship other than (primary religion houses of worship)?**

Appendix D: Data Confidence

The summary table presented below serves as a foundational reference for evaluating the confidence level of data used in Joshua Project's database updates, particularly regarding population, language, Christian adherents, evangelical presence, and location data. By categorizing sources into High, Medium, Low, and Unknown confidence levels, this table highlights the criteria that inform our assessment of reliability and verification. The goal is to enable thoughtful data integration while allowing for iterative improvements through team collaboration and field input. Additionally, age-based indicators will track the relevance of data over time, ensuring clarity on its timeliness and reliability. This framework is a starting point and open to refinement as we gain feedback and insights.

Confidence Level	**Key Characteristics**	**Source Type**	**Verification**
High	Specific, verified data; trusted source; transparent; 95% CI and 0.5% margin of error for survey sample	Specific quantitative data, census reports with direct data, random survey with low margin of error.	Cross-verified
Medium	Estimates or partial data; moderately reliable source; 95% CI and 1% margin of error	Extrapolated quantitative data, census data with extrapolation, random survey with low-moderate margin of error, Delphi Method with high level of consensus from	Limited verification
Low	Rough estimates; anecdotal evidence; informal source	Estimates based on broad observations, outdated or unreliable census data, Delphi method averaged with moderate consensus	No verification
Unknown	No data or unverifiable source	Unknown	No verification