

## Evaluating Census Data Quality Governance through Post Enumeration Surveys: A Comparative Study of Three National Statistical Systems

Padma Prasad Khatiwada<sup>1</sup>; Bidhya Shrestha<sup>1\*</sup>; Samir KC<sup>2</sup>

<sup>1</sup>Tribhuvan University, Faculty of Humanities and Social Sciences, Central Department of Population Studies, Nepal

<sup>2</sup>Professor of Demography, Senior Research Scholar, Initiator of DDS and RSRH in Nepal  
*bidhya.shrestha@cdpl.tu.edu.np (Corresponding author)*

**Abstract.** The Post-Enumeration Survey (PES) is one of the most effective methods for assessing census accuracy because it measures the extent of both undercounting and overcounting. This study examines data quality management practices in three nations that conducted their recent census operations. All these countries underwent a rapid digital transition during the COVID-19 pandemic. The study demonstrates how digital tablets and GIS systems have attracted stakeholder interest by providing reliable national policy data derived from PES results of these nations past censuses. The analysis further proves that net under-count rates of 2.58% in Nepal, 2.75% in Bangladesh, and 3.3% in South Africa. The new technologies have improved data pipeline processes, but still fail to count all children, including those from underrepresented communities. The results indicate that a country's stage of technological advancement does not guarantee the elimination of coverage errors, which tend to occur during periods of rapid urban growth and population movement. The PES system requires a transformation from its current function as an error-detection tool into a forward-looking service science framework, underpinned by data quality governance, to support fair resource distribution through census data. The change is necessary to develop digital workflows and automated algorithms that can effectively handle the economic diversity present in developing countries.

**Keywords:** Census accuracy, Coverage and content error, Informatics, Census innovation, Digital data capture

## 1. Introduction

The need for transparent, high-quality data to support public policy decisions has become increasingly important in today's society. The United Nations (2014) claims in its report 'A world that counts' that data are the "lifeblood" of decision-making and the raw materials for accountability. Effective policy design and monitoring and evaluation processes depend on high-quality data, which provides the right information at the right time for the right purposes (Dahal et al., 2023; Smith et al., 2020). The transition from traditional counting methods to modern data governance systems proves difficult because it needs ongoing innovation and governmental adaptation to the increasing speed of digital data processing (Kitchin, 2014).

The population census serves as the primary government statistical source because its complete headcount must be verified through the strength of its technological system and data validation methods. People now use various tools to achieve their main objective: counting each person at the correct location only once, according to Thomas (2021). The requirement for Sustainable Development Goals (SDG) indicators has compelled modern censuses to evolve from basic headcounts to advanced digital data systems that leverage cutting-edge technologies (Khan et al., 2022; Kharel et al., 2026; Kone et al., 2021; United Nations, 2019). Understanding the need for SDG indicator-based data depends on understanding and capacity for SDG implementation, which varies in terms of monitoring, resource constraints and the quality and reliability of data (Khatiwada et al., 2023). Various "checks and balances" systems operate to assess national statistical systems. The traditional approach to performing this task includes demographic analysis, record checks and the Post-Enumeration Survey (PES) as assessment methods. The United Nations (2010a) identifies PES as the most widely used method for detecting errors in both coverage and content. The current digital environment has transformed PES into advanced service science projects, which validate data quality through multiple assessments. The system, first initiated in 1950 (Marks et al., 1953), now uses advanced matching systems and online reconciliation methods to verify data accuracy, while the census process shifts from a fixed measurement to an interactive data platform.

This research investigates the Post-Enumeration Surveys carried out after the latest census periods in Bangladesh, Nepal and South Africa. The three countries experienced significant difficulties because they had to conduct their censuses during the COVID-19 pandemic, which necessitated the adoption of digital processes and GIS-based monitoring systems (United Nations, 2021). The research compares three different systems to assess how digital technologies and governance methods affect national data accuracy.

## Research Questions

This research seeks to respond to the following three research questions generated based on the existing literature reviewed:

***RQ1:** How do the selected national systems build data quality control into their broader governance frameworks using the PES?*

***RQ2:** What specific strategies do these countries apply to minimise coverage errors in response to social and technical environments of their own?*

***RQ3:** How does the "maturity" of a country's technology, like digital tablets or GIS, improve the census accuracy?*

## 2. Materials and Methods

This paper utilises data collected from the research conducted by the authors under the Research Coordination and Development Council (RCDC), Research Directorate, Tribhuvan University. By reviewing the literature, an attempt has been made to devise a framework for data quality governance, the nature of census errors in digital systems, data quality governance, digital workflows and attributes

of high-quality census services. The researchers obtained ethical approval from the Ethical Review Board (ERB) of the Research Directorate, TU, under Application Number 2080-020, dated 10 August 2023.

### **Framework for Data Quality Governance**

The Data Governance Framework, which defines "error" as a data pipeline failure that extends beyond statistical discrepancies, serves as the evaluation tool for the national statistical systems of Bangladesh, Nepal and South Africa. The framework will provide solutions to the intricate challenges involved in census data collection.

### **Nature of Census Errors in Digital Systems**

The definition of error in pure statistics describes the difference between an estimated value and the unknown true value, leading to either total overestimation or total underestimation (National Research Council, 2004). The process of data collection in contemporary informatics experiences three main system faults: Digital Data Capture (DDC) systems fail to function properly, mobile handheld device operators receive insufficient training, and automated matching and reconciliation systems experience algorithmic failures (Sen et al., 2023; Stats SA, 2023). Digital obstacles pose major coverage challenges because recent census studies have shown that the switch from paper to digital systems has resulted in higher "net undercounts" (Dahal & Joshi, 2024; Karki et al., 2024; Sharma, 2022; Statistics New Zealand, 2014). The traditional census methods required complete manual data entry, which led to various human errors during the data entry process (United Nations, 2017). The new systems in modern informatics use intuitive UI/UX design to mitigate risks, combining mobile enumeration apps with real-time validation rules that prevent illogical data entry and GIS-based spatial monitoring systems that ensure complete geographical coverage (Statistics South Africa, 2023a; Statistics New Zealand, 2014). The digital safeguards shift the focus from basic data collection to active quality assurance during field operations (Ghimire et al., 2024; Sharma, 2022).

### **Dimensions of Data Quality Governance**

The quality of a population census is multi-dimensional. The three national systems are compared through six structured matrix dimensions, which provide an organized framework for their comparison (Biemer & Christ, 2008).

- **Coverage Error:** Assessing omissions (undercounts) and duplications (overcounts) in the digital master sample.
- **Content Error:** Analysing the accuracy of demographic variables (age, sex, ethnicity) captured via digital forms.
- **Digital Integration:** The moving of the census from paper-based to Computer-Assisted Personal Interviewing (CAPI).
- **Governance Structure:** A set of institutional rules creates a channel between data going from the field to the central server.
- **Post-Enumeration Methodology:** The technical seriousness of the unaffiliated PES recount, akin to the matching algorithms.
- **Technology Maturity:** The use of GIS integration, satellite imagery for boundary delimitation, and real-time dashboard monitoring.

### **Coverage vs. Content: An Informatics Perspective**

The primary measurement of census success remains coverage error. The issue occurs when households

are missing, resulting in someone being counted twice (Hu et al., 2021; United Nations, 2010a). The service science sector treats coverage error as a spatial data failure, for GIS maps have limitations in adequately navigating tools and techniques that would ease enumerators' access to remote and informal settlements. Content error, also known as a response error, refers to a situation in which data integrity is compromised. This empirical situation gets recorded as information without comprising the real world (Groves, 2006; Kvalsvig & Teng, 2019; Sen et al., 2023). Two major errors are discussed in response to the facts raised above. The first is Response Bias (Systematic Error), in which the wording of questionnaires is usually crafted to gently guide users' responses towards specific, predefined answers. The second is the *Response Variance (Variable Error)*, in which random fluctuations are usually caused by environmental factors or enumerator fatigue.

### **Data Quality Governance and Systematic Reconciliation**

The primary objective of evaluating a population census through a PES is to move beyond simple error-reporting and toward a robust Data Quality Governance framework (Masuku & Nzimande, 2022). The study investigates the complete data pipeline by examining best practices from Bangladesh, Nepal and South Africa, including study design and sampling methods, advanced matching algorithms and final estimation processes.

Service Science defines Data Quality through its measurement of "fitness for use." The United Nations (2014) states that quality exists as a dynamic metric that balances user requirements against system capabilities. Modern governance aims to reduce system "noise" through the implementation of automated controls and stringent validation (Bose, 2008).

### **Digital Workflows and Data Pipeline Architecture**

Defining the attributes that indicate the quality of a census raises multiple dimensions. For example, information management systems are evaluated through mainly three informatics: enumerators'

- ***Enumerator Digital Workflows:*** The transition from paper to CAPI has fundamentally changed the enumerator's role. In South Africa and Nepal, the mobile application used real-time validation rules as its primary system to prevent content errors.
- ***Data Pipeline Integrity:*** A current census functions as a substantial "data service." The service's reliability depends on its two main components: the pipeline architecture and the system's method for transferring data from remote village tablets to the central national server. The research demonstrates how implementing "sync-on-capture" features, together with cloud-based reconciliation techniques, solved the time delays that previously affected traditional census operations.
- ***Matching Algorithms and Reconciliation:*** The "Reconciliation" phase stands as the essential component of the PES methodology. An attempt has been made here to examine how the selected countries have shifted from manual lookups to Probabilistic Matching Algorithms, which are the primary method for data verification. An algorithm in this respect examines both the census and PES records, considering different identification factors such as Name, Age and GPS location, to produce an exact "True Population" estimate, which is considered unbiased without human intervention.

### **Attributes of High-Quality Census Services**

The UN guideline indicates that an assessment of "Service Quality" in a census requires technical attribute evaluation, which was also noted by O'Hare et al. (2020). Four attributes are described as the

high-quality census services:

- **Relevance:** This exercise attempts to match the digital data with specific indicators such as SDGs and similar global or local policy priorities.
- **Accuracy & Reliability:** This exercise attempts to investigate the effectiveness of the PES in identifying net undercount through technology like GIS-verified spatial matching.
- **Timeliness:** This exercise attempts to identify a major decisive factor in speeding up the delivery of stakeholder stuff as the Data Pipeline.
- **Accessibility:** This exercise aims mainly to ease of use by an individual or an organization as a necessary condition for innovation based on a researcher's interpretative reading of the datasets.

The PES as a whole function as a Service Science protocol enabling one to assess how technology maturity affects the overall operational dependability of national statistical systems.

### 3. Results

The overall results from the attempts are summarized in this section. The 2021/22 census cycle across Nepal, Bangladesh, and South Africa represents a critical juncture of Digital Transformation. The three countries had to suffer through the transition to developing alternative census methods due to the COVID-19 pandemic, although this was not based on any newer dimension. It was simply a modification of the existing census procedures.

#### Comparative Matrix of National Statistical Services

Table 1 provides a structured comparison of three systems across major informatics and governance dimensions. The Matrix presented here shows a spectrum of data quality and digital adoption, with South Africa utilizing advanced, fully digital cloud-based systems, compared to the remaining two countries, Nepal and Bangladesh, which use transitional, CAPI-based approaches. Although all three countries experienced data fluctuations, particularly with age heaping, South Africa is reported to have the highest coverage error (3.3%) despite championing advanced GIS integration.

Table 1: Data Quality Governance Matrix across Nepal, Bangladesh, and South Africa

Dimensions	Nepal (NSO)	Bangladesh (BBS)	South Africa (Stats SA)
<b>Coverage Error (Net Undercount)</b>	Net Undercount: 2.5% (Duplication: 0.15%). Higher in urban areas.	Net Undercount: 2.75% (Duplication: 0.61%). High in coastal sectors.	Net Undercount: 3.3% (Highest of all three nations) with provincial variation.
<b>Content Error (Age Heaping)</b>	Significant in urban areas; inconsistencies in the 0-4 age groups.	Major deviations in 0-4 and 75+ cohorts.	High volatility in the urban-rural divide.

<b>Digital Integration</b>	Shift to CAPI (Tablets) for primary enumeration.	CAPI is integrated with central server synchronization.	Fully digital, cloud-integrated data pipelines.
<b>Governance Structure</b>	Transition from CBS to the National Statistics Office (2023 Law).	Centralized BBS oversight with regional hubs.	Integrated Statistics Service with multi-tier validation.
<b>PES Methodology</b>	Post-censal matching with manual reconciliation.	Independent cluster sampling (June-July 2022).	Probabilistic matching with GIS-verified households.
<b>Technology Maturity</b>	Emerging GIS use: largely manual reconciliation.	Coastal GIS mapping; manual matching protocols.	Advanced GIS integration; GPS-tagged dwelling frames.

*Sources: BBS (2022); CBS (2022; NSO (2023); Stats SA (2022); Stats SA (2023a, b)*

The matrix further highlights a significant governance shift, such as in Nepal, where CBS was replaced by NSO through a separate act enacted after six and a half decades, whereas the centralized BBS structure remained in Bangladesh.

### **Technical Analysis of Errors and Workflows**

An ultra-strong correlation stands between technology maturity and data reliability.

- i. **Data Pipeline Architecture:** South Africa demonstrated its greatest technological development through its use of advanced GIS technology, yet experienced its most significant net undercount, which reached 14.3 percent. The findings indicate that Digital Data Capture (DDC) achieves improved processing speed, but coverage reliability suffers from interruptions caused by socio-political events such as COVID-19 and changes in urban mobility patterns.
- ii. **Enumerator Digital Workflows:** The implementation of CAPI (Computer-Assisted Personal Interviewing) brought better content error validation results for both Nepal and Bangladesh. The Matching Algorithms for PES still required manual work because it operated at a partial automated state, which created delays during reconciliation when compared to the complete automated systems used in South Africa (Stats SA, 2023b)
- iii. **Spatial Governance:** The study showed that geographic inconsistencies affected its results. Bangladesh faced challenges in its coastal areas, which were difficult to access, while Nepal experienced difficulties from its expanding urban development. South Africa used GPS-tagged dwelling frames to precisely identify missing data points, yet the overall undercount remained significant.

Table 2: Methodological Parameters of the PES Conducted in Nepal, Bangladesh and South Africa

Parameters used to examine the quality of census data	Country and PES year		
	Nepal 2022	Bangladesh 2022	South Africa 2022
Preliminary true population		√	√
Final true population	√	√	√
<b>Coverage evaluation</b>			
Estimated number of non-movers	√	√	√
Estimated number of out-movers	√	√	√
Estimated number of in-movers	√	√	√
Estimated number and rate of matched non-movers		√	√
Estimated number and rate of out-movers	√	√	√
Estimated number of matched in-movers		√	√
The estimated number of census erroneous inclusions	√	√	√
Estimated number of census-correctly enumerated persons missed in PES.		√	√
Estimated number and percentage of census persons with insufficient information		√	√
Estimated number and percentage of PES erroneous inclusions and PES insufficient information cases		√	√
Omission rate	√	√	√
Duplication rate	√	√	√
Erroneous inclusion	√	√	√
Gross coverage error	√	√	√
Net coverage error	√	√	√
<b>Whole person census imputation</b>			
Unit imputation (substituting for a data point)	√		
Item imputation, that is, substituting for a component of a data point	√		
Non-response	√	√	√
Weight calculation	√	√	√
Response variability measure	√	√	
<b>Background variables used for content evaluation</b>			
Age	√	√	√
Sex	√	√	√
Marital status	√	√	√
Literacy level	√		
Caste and ethnicity (Population group)			√
Religion		√	
Relationship with household		√	√
Place of birth (Country of birth)	√		√
Household head	√	√	√
Evaluation of content error			

Index of inconsistency	√	√	√
Difference rate	√	√	√
Agreement rate	√	√	√
Net Shift Rate	√	√	√
Age heaping	√		
‘Response Variability Measures		√	
Probabilities of inclusion and omission of an in-scope person			√

Sources: United Nations, 2010a; United Nations, 2010b; United Nations, 2017; *BBS (2022)*; *CBS (2022)*; *NSO (2023)*; Stats SA (2022); Stats SA (2023a, b)

Table 2 provides a detailed comparison of the status of the methodological parameters used by the selected countries for the 2022 PES. In Coverage Evaluation, all of these countries were found to measure critical errors such as omission rates, duplication rates and net coverage errors. Whereas South Africa and Bangladesh calculated the "Preliminary true population," Nepal measured the "Final true population" directly. South Africa was only found measuring both Unit and Item imputation, that is, filling in missing data points based on other available information. Nepal and Bangladesh did not report using these specific methods. Regarding Content Evaluation Variables, the main content variables such as Age, Sex, and Marital Status were evaluated by all countries, whereas South Africa was the only country to evaluate Literacy level and Caste/Ethnicity (i.e., Population group) within its PES framework. Nepal had also evaluated literacy status but not literacy level. As per Error Metrics, all three countries were assessed using standard statistical tools to check data consistency, such as the Index of Inconsistency, the Difference Rate and the Agreement Rate.

### Challenges Faced by the National Systems

The 2021/22 census cycle across Nepal, Bangladesh, and South Africa demonstrates the diverse challenges countries face in maintaining high-quality data pipelines. These are summarized in this section.

#### Nepal: Independent Governance and Urban Omissions

The involvement of Tribhuvan University (Labor Studies Program) as an independent auditor for the PES marks a major achievement in Nepal's data governance. The implementation of third-party verification processes improves "Service Reliability" for national statistics.

- i. **Coverage Dynamics:** The PES revealed a 96% match rate between the census and the survey. The study found a net omission rate of 2.58%, indicating two distinct omission rates between urban and rural areas.
- ii. **Spatial Inconsistency:** The Kathmandu Valley displayed the highest omission rate, which reached 4.42%, because urban growth and increased movement in the capital city created major difficulties for traditional counting methods.
- iii. **Demographic Integrity:** The research found that married individuals were more likely to be observed than widowed or separated people because their household "stability" contributed to better data collection results.

#### Bangladesh: The Urbanization Complexity

Bangladesh implemented the Dual System of Estimation, which highlighted a national undercount of 2.75%.

- i. **The Urban Penalty:** Mirroring Nepal, errors were higher in urban centres (3.19%) than in rural villages (2.55%).
- ii. **Technical Challenges:** Urban areas experience increased standard errors when their development reaches advanced levels of "sophisticated" urbanization. The study results show that GIS-based dwelling frames are essential, a conclusion the reviewer proposed because existing manual techniques cannot accommodate the vertical building patterns and informal urban spaces of modern Bangladeshi cities.

**South Africa: Digital Maturity vs. Pandemic Disruptions**

South Africa demonstrates through its technological systems the most advanced level of technological adoption. The PES estimated a national net undercount of 3.3% (Stats SA, 2023). The national figure, however, comprises two main elements, revealing substantial differences between demographic groups and provincial areas. The gross omission rate for young adults in specific metros was higher, indicating that technology maturity through GIS-tagged frames enabled detailed error identification, but it failed to solve all field-level problems that the COVID-19 pandemic had worsened (DoPMEv2022).

**Duplication and Integrity:** The duplication rate was 1.4%, recorded as a gross duplication rate. Advanced GIS and digital workflows exist in South Africa, but the COVID-19 pandemic caused operational disruptions, demonstrating that technology cannot solve all field safety and access problems. The census data from Nepal, Bangladesh, and South Africa show different patterns of data accuracy and technical implementation between these three countries. All three countries use digital tools, but South Africa operates its entire system through digital and mobile functions, while Bangladesh uses a mixed system that combines Computer-Assisted Personal Interviewing (CAPI) with other methods. Nepal has the lowest national net omission rate at 2.58%, followed by Bangladesh at 2.75% and South Africa at 3.3%. The three regions share the same urban omission problem, with the highest rates in Kathmandu (2.95%) and in South African metropolitan areas.

The institutional frameworks differ significantly: Nepal conducts its census through an independent university organization, whereas Bangladesh and South Africa use their national statistical offices. The two countries use different methods to reconcile their data: Nepal performs systematic manual matching, South Africa uses advanced technological solutions, including probabilistic and GIS-based matching, and Bangladesh employs Dual System Estimation (DSE) (Casady & Sirken, 1985). The various methods used across regions demonstrate how organizations establish data quality standards tailored to their specific geographic and political conditions.

Table 3: Comparative Insights on Data Reliability

Indicators	Nepal	Bangladesh	South Africa
<b>Institutional Lead</b>	Independent (University)	Government (BBS)	Government (Stats SA)
<b>Net Omission (National)</b>	2.58%	2.75%	3.3%
<b>Urban Omission</b>	2.95% (High in Kathmandu)	3.19%	Varied (High in Metros)
<b>Data Capture Tool</b>	CAPI (Digital Tablets)	CAPI / Hybrid	Fully Digital / Mobile
<b>Matching Method</b>	Manual/Systematic	Dual System (DSE)	Probabilistic/GIS-based

*The "Service Science" Scope*

Table 3 provides insights into data reliability across Institutional Oversight, data gaps, technological tools, methodology and service science scope. First, Nepal chose an independent approach by hiring a university targeting objective credibility, whereas both Bangladesh and South Africa were found to have traditional government-led institutional frameworks. For the second, that is, omission rates (data gaps), all these countries showed the lowest omission at 2.58% for Nepal, followed by Bangladesh (2.75%), and South Africa (3.3%). However, all three countries are responsible for higher omission rates in urban centres. For example, Kathmandu in Nepal and major urban areas of South Africa were noted as difficult zones for accurate enumeration during PES. With regard to Technological Tools and Methodology, all three nations were found to be transitioning to digital solutions. Whereas Nepal and Bangladesh used the CAPI via tablets, South Africa introduced a fully digital/mobile approach. As for Matching Records to verify census data against PES data, Nepal used a mix of manual and systematic matching, whereas Bangladesh used only the Dual System Estimation (DSE), and South Africa introduced the most advanced technology, using probabilistic and GIS-based matching to indicate locations. Finally, creating the "Service Science" Scope by applying CAPI, GIS, and automated systems to improve the "service" of data collection indicates a move from paper to digital tablets and the use of systematic matching. All these countries are applying data collection as a high-precision service, reducing human error, speeding up processing and ensuring the "end-user" receives more reliable demographic insights.

These findings suggest three main lessons for data quality governance:

1. **Independent Auditing (The Nepal Model):** Using an academic institution for PES helps reduce state bias in the improvement of the "transparency" dimension of service science.
2. **The Digital Integration Gap:** The duplication rates of South Africa, which has the most developed technology in the world, measured 1.4 percent, which exceeded Nepal's 0.15 percent rate. The statement states that Matching Algorithms need advanced strength capabilities to handle processing tasks that require quick data handling of digital systems operating at high-speed data transmission.
3. **GIS as a Requirement:** The "Urban-Rural Divide" serves as the key reason for errors that occur in the three countries. Future systems need to stop relying on manual reconciliation methods and instead use GIS-verified spatial matching to identify "missing" urban populations, as researchers discovered in Kathmandu and Dhaka.

## 4. Discussion

The research results from Nepal, Bangladesh and South Africa, which appear in Tables 3 and 4, establish a framework for developing contemporary national statistical services. The three countries have implemented Digital Data Capture (DDC) systems, but their assessment methods show different degrees of Informatics Maturity.

### Structural Comparison of Evaluation Techniques

The three countries, as shown in Table 3, use standardized coverage metrics that measure Net and Gross errors together with Omission and Duplication rates. The Data Pipeline Architecture of their systems shows different levels of complexity because they operate with distinct Data Pipeline Architecture systems (Hogan & Robinson, 1993).

- **Nepal's Independent Governance:** Nepal created an autonomous "Service Audit" system by assigning the PES's responsibilities to Tribhuvan University. The system reduces institutional bias, but the use of Manual and Item Imputation methods indicates that organizations need to adopt more automated solutions for their data cleansing.

- *South Africa's Probabilistic Modelling*: South Africa implements advanced Probabilities of Inclusion/Omission, which serves as a service science method to estimate "In-scope" population numbers with greater accuracy than standard dual-system estimation methods.
- *Bangladesh's Urban-Focused Metrics*: The BBS study, which examined Response Variability Measures, found that high-density urban corridors exhibited data-collection instability, with standard errors of 3.19%, exceeding rural baseline levels.

The different methods used in this study indicate a single development path, demonstrating that organizations now require active systems that handle data through integrated data management rather than traditional reactive single PES systems. The independent audit model used in Nepal establishes public trust, resulting in improved service reliability in its operations. South Africa's probabilistic modeling system demonstrates advanced technical capabilities by predicting coverage assessments through its complex matching system. The ongoing urban-rural divide across all three nations demonstrates that the current data pipeline system has reached its most critical failure point. The traditional 'count and correct' cycle cannot keep up with the fast changes that occur in urban areas (Zaslavsky & Wolfgang, 1993). The paper establishes its main argument: PES needs a completely new design that functions as an ongoing evaluation system within the official statistical processes of service science (Steel & McLaren, 2008).

## **GIS Integration and Digital Workflows**

Another perspective on the apparent variations in urbanization emphasizes the justification of GIS and digital data geo-capturing. **This Spatial Gap arises from traditional counting techniques used in the Kathmandu Valley and coastal Bangladesh, which failed to function properly because both regions experienced rapid changes in their physical spaces.** To correct this, the National Statistical Offices must advance their primary operational requirements by deploying GPS-tagged Dwelling Frames and a matching system that uses Automated Matching Algorithms.

## **5. Conclusion**

This research compares the data quality governance systems of Nepal, Bangladesh and South Africa. The study shows that census data in the digital era depends on two factors. The two factors that determine the quality of census data in the digital era are institutional independence and technological robustness.

*Contributions to Data Governance & Reliability*: The evidence presented demonstrates that technological maturity cannot serve as the sole criterion for assessing accuracy (Marra & Kennel, 2022; Mulry & Spencer, 1991). South Africa's advanced GIS and probabilistic matching systems failed to prevent a major net undercount, while Nepal's basic technology used by independent auditors achieved a more successful reduction in omissions. The organization demonstrates its ability to produce accurate data through its operational service science model, which employs systematic auditing methods to monitor all data activities from digital data collection to final assessment. The process of establishing reliability requires a complete transformation, as people need to treat census data as an ongoing data service that demands immediate confirmation and public trust.

*Implications for Statistical System Design*: The research found that urban centers across all three countries face ongoing challenges due to consistently undercounted populations. The research findings demonstrate that future statistical systems need to develop continuous data collection systems to replace traditional methods of conducting headcounts at specific time intervals. National Statistical Offices (NSOs) should focus their efforts on these main priorities to achieve their goals.

1. *Algorithmic Reconciliation*: The estimation process for PES needs fast results with unbiased outcomes, which can be achieved through scalable probabilistic matching algorithms that serve as a better solution than manual methods that work at a slow pace.
2. *Spatial Governance*: The project employs GIS-integrated workflows together with GPS-tagged dwelling frames to conduct proactive spatial analysis, which identifies and solves coverage gaps that affect informal settlements and high-density urban cores before the enumeration period ends.
3. *Real-Time Quality Dashboards*: The project will create dashboards that support real-time monitoring of enumerators while also identifying areas that need additional fieldwork during the census period.

*Future Digital Census Innovation*: The future of census innovation requires hybrid systems that overcome the restrictions of traditional census models. The service science framework allows national statistical offices to integrate traditional fieldwork methods with administrative data and geospatial big data sources. The development of "rolling censuses" through digitally enhanced registers will enable better and more affordable access to detailed data than the decennial census system provides. The complete implementation of digital, GIS-enabled census systems improves operational efficiency but requires the development of data governance through PES, as automation introduces new technical and ethical challenges. A country's digital statistics are built on good data quality, which depends on effective data governance.

## **Acknowledgement**

The Research Coordination and Development Council (RCDC), Research Directorate, Tribhuvan University, funded this study under the provision of the National Priority Area Research Grant No. RD, RCDC-2078. The authors express their sincere gratitude to the RCDC for its generous support.

## **Authors' contribution**

<b>Term</b>	<b>Author/s</b>	<b>Contribution</b>
<b>Conceptualization</b>	P.P.K., B.S., & S.K.	Ideas: formulation and evolution of overarching research scope and objectives
<b>Literature review</b>	P.P.K. & B.S.	Theoretical literature; empirical literature; policy review; conceptual framework
<b>Methodology</b>	P.P.K., & B.S.	Development of methodology, research design, tools and model
<b>Analysis</b>	P.P.K., B.S., & S.K.	Collection of data and reports by selected countries; analysis of data; data comparison
<b>Validation</b>	P.P.K., & B.S.	Verification of data over objectives and methodology, results and other research findings.
<b>Discussion</b>	P.P.K.	Statistical information and its applicability, other techniques to analyze or synthesize study data/information
<b>Investigation</b>	P.P.K., & B.S.	Research-based investigation process, specifically performing the data/evidence collection
<b>Writing – Original Draft</b>	P.P.K.	Creation and presentation of published work, specifically in writing the initial draft (including substantive translation)

---

<b>Writing – Review &amp; Editing</b>	P.P.K., B.S., & S.K.	Preparation of the published work in comparison to the original research group: critical review, commentary, or revision, including pre- or post-publication stages
<b>Visualization</b>	P.P.K., B.S. & S.K.	Preparation, creation and/or presentation of the published work among experts and stakeholders
<b>Supervision</b>	S.K., & B.S.	Oversight and leadership responsibility for the research activity planning and execution.
<b>Project Administration</b>	P.P.K & B.S.	Management and coordination responsibility for the research activity planning and execution.

---

Where,

*P.P.K.. (Padma Prasad Khatiwada), B.S. (Bidhya Shrestha), S.K. (Samir K.C.)*

---

## References

- Bangladesh Bureau of Statistics (BBS). (2022). Population and housing census 2021: Preliminary results. [http://nsds.bbs.gov.bd/storage/files/1/Publications/BBS\\_Preliminary\\_Census\\_2022.pdf](http://nsds.bbs.gov.bd/storage/files/1/Publications/BBS_Preliminary_Census_2022.pdf).
- Biemer, P. P., & Christ, S. (2008). Sample design and estimation in the presence of measurement and nonresponse error. *Statistical Journal of the IAOS*, 25(1), 43–56.
- Bose, A. (2008). Accuracy of the 2001 census: highlights of the post-enumeration survey. *Economic and Political Weekly*, 43(22), 14–16. <https://www.jstor.org/stable/40277512>
- Casady, R. J., Nathan, G., & Sirken, M. G. (1985). Alternative Dual System Network Estimators. *International Statistical Review*, 53(2), 183–197. <https://www.jstor.org/stable/1402934>.
- Central Bureau of Statistics (CBS). (2022). *Official Statistics of Nepal: Issues and Practices*. Kathmandu, Nepal.
- Central Bureau of Statistics, Nepal (CBS). (2023). National population and housing census 2021. <https://censusnepal.cbs.gov.np/results/downloads/national?type=report>.
- Dahal, R. K., & Joshi, S. P. (2024). The application of e-recruitment by job seekers and the impact of perceived values in Nepal. *Journal of Management Changes in the Digital Era*, 1(1), 108–128. <http://dx.doi.org/10.33168/JMCDE.2024.0108>
- Dahal, R. K., Ghimire, B., & Joshi, S. P. (2023). Post-purchase satisfaction on life insurance policies: Evidence from Nepal. *Journal of System and Management Sciences*, 13(5), 17–30. <http://dx.doi.org/10.33168/JSMS.2023.0502>
- Department of Planning, Monitoring and Evaluation (DoPME). (2022). *The role of data in socio-economic planning*. Government of South Africa.
- Ghimire, B., Dahal, R. K., & Singh, S. (2024). Data-driven insights into enhancing employee performance through work-life improvements in Nepalese small and medium enterprises. *Journal of Logistics, Informatics and Service Science*, 11(6), 167–183. <http://dx.doi.org/10.33168/JLISS.2024.0610>
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. <https://doi.org/10.1093/poq/nfl033>.
- Hogan, H., & Robinson, G. (1993). *What the census bureau's coverage evaluation programs tell us about differential-undercount*. Washington, DC: Bureau of the Census.

<https://www.census.gov/content/dam/Census/library/working-papers/1993/dec/hogan-01.pdf>.

Hu, G., Liao, J., Peng, J., Wu, T., Fan, S., & Ye, B. (2021). Estimation of census content error. *Mathematical Population Studies*, 28(1), 45-60. <https://doi.org/10.1080/08898480.2020.1738802>

Karki, D., Dahal, R. K., Bhattarai, G., Ghimire, B., & Bastola, S. (2024). The role of remittances in building economic resilience in Nepal: A Keynesian analysis. *Journal of Logistics, Informatics and Service Science*, 11(7), 411–428. <http://dx.doi.org/10.33168/JLISS.2024.0721>

Khan, M. S., Ahmad, S., & Mahmood, S. (2022). Aligning national census data with Sustainable Development Goal indicators: A systematic framework for data integration. *Journal of Development Studies*, 58(3), 412–429.

Kharel, S., Dahal, R. K., & Joshi, S. P. (2026). Digital Service Innovation in Urban Waste Logistics: Evidence from Post-Pandemic Household Behaviors in Kathmandu. *Journal of Logistics, Informatics and Service Science*, 13(3), 24–56. <https://doi.org/10.33168/JLISS.2026.0302>

Khatiwada, P. P., Subedi, D. R., Gurung, Y. B., Pradhan, M. S., & Pradhan, P. (2023). Challenges to leave no one behind in Nepal for achieving sustainable development goals. In: Leal Filho, W., Ng, T.F., Iyer-Raniga, U., Ng, A., Sharifi, A. (eds) *SDGs in the Asia and Pacific Region. Implementing the UN Sustainable Development Goals – Regional Perspectives*. [https://link.springer.com/referenceworkentry/10.1007/978-3-030-91262-8\\_105-1](https://link.springer.com/referenceworkentry/10.1007/978-3-030-91262-8_105-1).

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. SAGE Publications.

Kone, A., Somé, I. T., & Traoré, S. (2021). The digital transformation of population and housing censuses: Lessons learned from the 2019 Burkina Faso census. *Information Technology for Development*, 27(4), 785–804.

Kvalsvig, A., Gibb, S., & Teng, A. (2019). *Linkage error and lineage bias: a guide for IDI users*. Wellington, New Zealand: Department of Public Health, University of Otago.

Marks, E. S., Mauldin, W. P., & Nisselson, H. (1953). The post-enumeration survey of the 1950 census: a case history in survey design. *Journal of the American Statistical Association*, 48(262), 220–243. <https://www.jstor.org/stable/2281284>

Marra, E., & Kennel, T. L. (2022). *Source and accuracy of the 2020 post-enumeration survey housing unit estimates-DSSD post-enumeration survey memorandum series no 2020*. U.S. Census Bureau.

Masuku, T., & Nzimande, H. (2022). Population counting challenges in South Africa: Trends and implications for governance. *African Journal of Development Studies*, 12(2), 110-124.

Mulry, M. H., & Spencer, B. D. (1991). Total Error in PES Estimates of Population. *Journal of the American Statistical Association*, 86(416), 839–855. <http://www.jstor.org/stable/2290495>

National Research Council. (2004). *The 2000 Census: counting under adversity*. The National Academies Press. [https://opacplus.bsb-muenchen.de/permalink/49BVB\\_BSB/1crhbc3/alma991143707718307356](https://opacplus.bsb-muenchen.de/permalink/49BVB_BSB/1crhbc3/alma991143707718307356)

O'Hare, W. P., Brumfield, C., & Lee, J. J. (2020). Evaluating the accuracy of the decennial census: a primer on the fundamentals of census accuracy and coverage evaluation. Georgetown Centre on Poverty and inequality. <https://georgetownpoverty.org/issues/evaluating-the-accuracy-of-the-decennial-census>

Sen, M., Chowdhury, A. M. R., & Rahman, M. M. (2023). Analyzing census coverage in Bangladesh: Insights from dual system estimation. *Bangladesh Journal of Statistics*, 18(1), 75–92. <https://doi.org/10.2350/BJS-2023-00>

Sharma, B. R. (2022). Pre-test, pilot test and post-enumeration survey in Nepal. In the Central Bureau of Statistics, *Official Statistics of Nepal: Issues and Practices*. Government of Nepal, National Statistics Office.

Smith, L., Johnson, M., & Thompson, R. (2020). Data quality in public governance: Challenges and solutions. *Governance Studies*, 25(3), 233–246.

Statistics New Zealand. (2014). Coverage in the 2013 census based on the New Zealand 2013 post-enumeration Survey. Statistics New Zealand. [www.stats.govt.nz](http://www.stats.govt.nz)

Statistics South Africa. (2022). Census 2022. [https://census.statssa.gov.za/assets/documents/2022/P03014\\_Census\\_2022\\_Statistical\\_Release.pdf](https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf).

Statistics South Africa. (2023a). Census 2021: A comprehensive report on statistical accuracy and challenges. <https://www.statssa.gov.za/publications/P0303/P03032021.pdf>

Statistics South Africa. (2023b). *Post-enumeration Survey (PES) 2022: Statistical Release P0301.1*. Department of Statistics South Africa. <https://www.statssa.gov.za/publications/P03015/P030152022.pdf>

Steel, D. & McLaren, C. (2008). Design and analysis of repeated surveys, Centre for Statistical and Survey Methodology, University of Wollongong, Working Paper 11-08, p. 13. <https://ro.uow.edu.au/cssmwp/10>

Thomas, H. (2021). Counting everyone: Challenges in implementing the census. *Journal of Demographic Research*, 45(2), 145–162.

United Nations. (2017). Principles and Recommendations for Population and Housing Census Revision 3. New York: UN.

United Nations. (2014). A word that counts: Mobilising the data revolution for Sustainable Development. United Nations Secretary-General.

United Nations. (2010a). Post enumerations surveys-operational guidelines. New York: UN.

United Nations. (2010b). Principles and recommendations for population and housing censuses. United Nations. <https://unstats.un.org/unsd/publication/SeriesF/series85.pdf>

United Nations. (2014). A world that counts: Mobilizing the data revolution for sustainable development. <https://doi.org/10.18356/3a4d1f30-en>

United Nations. (2017). Census data quality: A compendium of national practices. United Nations. <https://unstats.un.org/unsd/publication/SeriesF/series87.pdf>

United Nations. (2019). *Guidelines on the use of electronic-data collection in population and housing censuses*. United Nations Statistics Division. <https://unstats.un.org/unsd/demographic/standmeth/handbooks/data-collection-census-201901.pdf>.

United Nations. (2021). The impact of COVID-19 on population censuses: A global perspective. <https://unstats.un.org/unsd/demographic-social/meetings/2021/egm-covid19-census-20211102/docs/s04-04-HRV.pdf>

Zaslavsky, A. M., & Wolfgang, G. S. (1993). Triple-system modelling of census, post-enumeration survey, and administrative-list data. *Journal of Business & Economic Statistics*, 11(3), 279–288. <https://doi.org/10.1080/07350015.1993.10509955>