

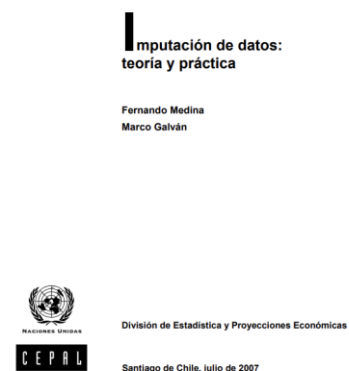
Metodologías de imputación aplicadas en encuestas de Hogares y Censos de población

Imputation methodologies applied on
household surveys and population census

País coordinador: Chile

Antecedentes / Antecedents

- Medina y Galván (2007) desarrollaron una guía para CEPAL sobre imputación de datos en encuestas de hogares, analizando sus efectos en indicadores de pobreza y desigualdad.
- Revisión de técnicas de manejo de no respuesta en encuestas de ingresos y gastos 2018 – 2019.
- Medina and Galvan developed a guide for ECLAC on data imputation in household surveys, analyzing their effects on poverty and inequality indicator
- Review of non-response management techniques in income and expense surveys 2018 - 2019.



Avance / Progress

- Conformación del grupo de trabajo (15 países y CELADE)
 - Elaboración de índice comentado
 - Revisión por parte de los países del índice
 - Presentación IBGE experiencia de imputación en Censo 2022
 - Envío consulta diagnóstico acerca de prácticas de imputación en encuestas de hogares y censos de población
- Formation of the working group (15 countries and CELADE)
 - Elaboration of the proposed annotated index
 - Review by the countries of the index
 - IBGE presentation about Census 2022
 - Sending diagnostic consultation regarding practices of imputation in household surveys and population census



Índice propuesto / Proposed index

- El documento de “Metodologías de imputación aplicadas en encuestas de Hogares y Censos de Población” busca instaurar estándares regionales de buenas prácticas adaptables a los países de la región, en el marco de la Conferencia de Estadística de las Américas, que proporcione a los equipos técnicos regionales los lineamientos necesarios para abordar de manera efectiva la imputación de datos faltantes en las encuestas de hogares y censos de población.
- Este manual debe servir como referencia práctica y teórica para mejorar la precisión y confiabilidad de los datos estadísticos generados, recopilando la experiencia regional en materia de imputación.
- The document "Imputation Methodologies Applied in Household Surveys and Population Censuses" seeks to establish regional standards of good practices adaptable to the countries of the region, within the framework of the Statistical Conference of the Americas, which provide regional technical teams with the necessary guidelines to effectively address the imputation of missing data in household surveys and population censuses.
- This manual should serve as a practical and theoretical reference to improve the accuracy and reliability of the statistical data generated, compiling regional experience in imputation.

Índice propuesto / Index plan



ÍNDICE COMENTADO DEL DOCUMENTO: RECOMENDACIONES METODOLÓGICAS SOBRE IMPUTACIÓN EN CENSOS Y ENCUESTAS DE HOGARES

TABLA DE CONTENIDO

1.	Antecedentes	12.	Conclusiones
2.	Definición y Tipos de Datos Faltantes	12.1.	Resumen de Hallazgos
2.1.	Concepto de Datos Faltantes	12.2.	Reflexiones Finales
2.2.	Clasificación de Datos Faltantes	13.	Bibliografía
3.	Objetivos Teóricos y Prácticos de la Imputación	14.	Anexos
3.1.	Justificación de la Imputación	14.1.	Estadísticas Complementarias
3.2.	Impacto en la Calidad de los Datos	14.2.	Paquetes Estadísticos Recomendados
4.	Patrones y Distribución de los Datos Faltantes	14.3.	Ejemplos de Comandos
4.1.	Análisis de Patrones Comunes		
4.2.	Métodos para Detectar y Manejar Patrones		
5.	Métodos Tradicionales de Imputación		
5.1.	Listwise Deletion		
5.2.	Pairwise Deletion		
6.	Reponderación		
7.	Métodos Modernos de Imputación		
7.1.	Imputación Simple		
7.2.	Imputación Múltiple		
7.3.	Técnicas de Aprendizaje Automático		
8.	Procesos de Imputación en la ronda de censos 2020		
8.1.	Estrategia de Imputación		
8.2.	Implementación y Comparación		
8.3.	Recomendaciones y Buenas Prácticas		
9.	Imputación en Encuestas de hogares		
9.1.	Desafíos Específicos		
9.2.	Ejemplos de Implementación		
9.3.	Efectos de la Imputación en la Pobreza y la Desigualdad		
9.4.	Recomendaciones y Buenas Prácticas		
10.	Estudios de Caso y Aplicaciones Prácticas avanzadas		
11.	Consideraciones Prácticas y Recomendaciones		
11.1.	Recomendaciones de Implementación		
11.2.	Selección de Métodos		

Índice en discusión / index in progress

Exposiciones países/ Countries expositions

- Establecer un intercambio con las oficinas nacionales de estadística de la región que permita la transmisión de conocimientos, buenas prácticas y experiencias exitosas en el campo de la imputación de datos en encuestas de hogares y censos de población.
- Presentación IBGE acerca del proceso de imputación en Censo demográfico Brasileiro 2022
- Establish an exchange with the national statistical offices of the region to facilitate the transmission of knowledge, best practice and successful experiences in the field of data imputation in household surveys and population censuses
- IBGE presentation about imputation process in Census 2022

Censo Demográfico Brasileiro 2022

✓ O uso dos DMCs (Dispositivos Móveis de Coleta) desde o Censo Demográfico de 2010 permitiu incorporar críticas no momento da coleta de dados:

- Fluxo do questionário;
- Críticas de valores válidos;
- Críticas de consistência.



Imputação no Censo Demográfico 2022

Classe socioeconômica do setor

- ✓ Escore do domicílio: soma das pontuações das variáveis
- ✓ Escore mínimo 0 e máximo 6.
- ✓ Escore médio: média dos escores dos domicílios do setor.

- Classe 1: se o escore médio era inferior a 3;
- Classe 2: se escore estava entre 3 e 5 (exclusive);
- Classe 3: se estava entre 5 e 6 (exclusive);
- Classe 4: se o escore era 6.

✓ Quanto mais baixa a classe, piores as condições socioeconômicas do domicílio.

IBGE
Instituto Brasileiro de Geografia e Estatística

Consulta diagnóstico/ Diagnostic consultation

- Se envió una consulta diagnóstica a los países participantes para sistematizar las experiencias institucionales.
- Consulta por Encuesta y Censo
- Observar el panorama de las oficinas nacionales de estadística en cuanto a imputación en encuestas y censos.
- A diagnostic consultation was sent to the participating countries to systematize institutional experiences
- Observe the landscape of national statistical offices regarding imputation in household surveys and population census



Consulta diagnóstica sobre las prácticas de imputación aplicadas en encuestas de hogares y censos de población

En el marco de la Conferencia Estadística de las Américas de la Comisión Económica para América Latina y el Caribe (CEPAL), se han definido grupos de países para trabajar en las distintas propuestas a ser abordadas durante el bienio 2024-2025. Uno de los grupos, tiene como objetivo sistematizar la experiencia internacional, con foco en los organismos nacionales de estadística de la región, que permita elaborar recomendaciones para la implementación de metodologías de imputación de datos en encuestas de hogares y censos de población.

Para lograr este objetivo, se precisa conocer las prácticas y experiencias de los países en esta materia, abordando además, otras temáticas que guardan relación con el objetivo planteado para el grupo de trabajo, tales como aspectos asociados a la repercusión en el cálculo de la varianza y medición del error de muestreo por parte de los usuarios de las encuestas y censos de población.

En este sentido, mediante el siguiente formulario, se busca recopilar información sobre los criterios aplicados por las distintas Oficinas Nacionales de Estadística (ONE) para determinar las prácticas de imputación utilizadas en encuestas de hogares y censos de población. Se pretende conocer las metodologías empleadas para imputar datos faltantes, las técnicas y herramientas utilizadas para llevar a cabo esta imputación, y cómo estas prácticas afectan la calidad y fiabilidad de los datos recogidos. Además, se espera obtener información sobre los desafíos encontrados en la implementación de estas prácticas, y las experiencias específicas en el contexto de encuestas continuas de hogares y censos, con un enfoque particular en la imputación de datos relacionados con la población, el mercado laboral, los ingresos y los gastos.

El formulario está estructurado en tres partes. Una primera parte para recopilar la información de la ONE y de su correspondiente punto focal o informante. La segunda parte, contiene una serie de preguntas formuladas en forma general sobre los tópicos de interés, para concluir con una tercera parte, formulada mediante fichas que buscan recopilar información de los tópicos de interés, pero específica de cada estudio.

Con el fin de complementar y referenciar la información obtenida mediante este formulario, se solicita que sea informada la documentación que desarrolla con mayor detalle, alguno de los tópicos consultados, mediante su envío como material complementario. Para referenciar los documentos que se adjuntarán o informarán, se deben indicar los campos:

- Nombre del documento.
- Tema que documenta.
- Link (si está disponible).
- Apartado y páginas en el cual localizar la información requerida.

Para el envío de la respuesta, así como de las preguntas que puedan surgir durante el proceso, deben comunicarse con: fmolinaj@ine.gob.cl (INE-Chile), con copia a deanlopez@ine.gob.cl (INE-Chile), obarqueroa@ine.gob.cl (INE-Chile) y andres.gutierrez@cepal.org (Punto focal CEPAL).



II. CONSULTA POR CENSO DE POBLACIÓN Y VARIABLE IMPUTADA (FICHA)			
Ficha del estudio	Año de último censo realizado		
Si aplica igual metodología para un distinto grupo de variables	Porcentaje de datos faltantes	Discreta	Continua
menciones en una sola ficha	Criterios utilizados para decidir la imputación de la variable		

Métodos de imputación utilizados	Métodos de imputación que emplea según la variable de interés (Ejemplos: imputación por promedios, por regresión, vector más cercano, donación aleatoria, hot deck, etc.)	Tipo de imputación	Marque la casilla		¿Puede el usuario identificar si la variable fue imputada? Si/No
			Si	No	
		Por promedios/medias			
		Regresión			
		Donación aleatoria			
		Hot Deck			
		Árboles de decisión			
		Imputación por mediana			
		Imputación múltiple			
		Otro. Especifique			
	Señale las variables utilizadas en caso de usar información auxiliar				
	¿Utiliza registros administrativos para la aplicación de las metodologías?	Si	No		
		¿Cuántos?			

Criterios	Criterios utilizados para decidir la imputación de una variable	Tipo de imputación	Marque la casilla	Indique valor utilizado (Si aplica)
		Porcentaje de datos faltantes		

Próximos pasos / Next Step

- Compilación y análisis resultado diagnóstico
 - Discusión y aprobación final índice documento
 - Presentación de experiencias internacionales
 - Envío de avance de Documento para revisión
- Compilation and analysis of diagnostic results
 - Final discussion and approval of the document index
 - Presentation of international experiences



Vigesimotercera Reunión del Comité Ejecutivo de la
Conferencia Estadística de las Américas de la CEPAL
28 y 29 de agosto de 2024



Muchas gracias
Thank you

País coordinador: Chile