



**UNIVERSIDAD AUTÓNOMA  
DE AGUASCALIENTES**

**CENTRO DE CIENCIAS BÁSICAS**

**DEPARTAMENTO DE SISTEMAS DE INFORMACIÓN**

**TRABAJO PRÁCTICO**

**ANÁLISIS DEL PROCESO DE SUPERVISIÓN DEL PROYECTO “ENCUESTAS  
EN HOGARES” DEL INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA  
(INEGI) PARA CONTROL Y TOMA DE DECISIONES**

**PRESENTA**

**I.S.C. Juan Eduardo Rioja Romo**

**PARA OBTENER EL GRADO DE MAESTRO EN INFORMÁTICA Y  
TECNOLOGÍAS COMPUTACIONALES**

**TUTOR**

**Mtro. Jorge Eduardo Macías Luévano**

**COMITÉ TUTORAL**

**Dra. Lizeth Itziguery Solano Romo**

**Dr. Juan Muñoz López**

**Aguascalientes, Ags, 24 de junio de 2020**

CARTA DE VOTO APROBATORIO  
COMITÉ TUTORAL

M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ  
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

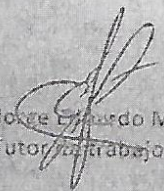
PRESENTE

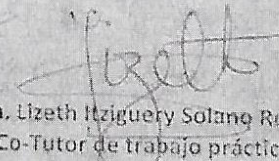
Por medio del presente como Miembros del Comité Tutorial designado del estudiante JUAN EDUARDO RIOJA ROMO con ID 15321 quien realizó el titulado: ANÁLISIS DEL PROCESO DE SUPERVISIÓN DEL PROYECTO "ENCUESTAS EN HOGARES" DEL INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA (INEGI) PARA CONTROL Y TOMA DE DECISIONES, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia damos nuestro consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que nos permitimos emitir el VOTO APROBATORIO, para que él pueda proceder a imprimirlo así como continuar con el procedimiento administrativo para la obtención del grado.


Ponemos lo anterior a su digna consideración y sin otro particular por el momento, le enviamos un cordial saludo.

ATENTAMENTE  
"Se Lumen Proferre"

Aguascalientes, Ags., a 23 de junio de 2020.

  
Mtro. Jorge Eduardo Macías Luévano  
Tutor de trabajo práctico

  
Dra. Lizeth Itziguery Solano Romo  
Co-Tutor de trabajo práctico

  
Dr. Juan Muñoz López  
Asesor de trabajo práctico

C.C.P. - Interesada  
C.C.P. - Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo a Posgrado  
Revisado por: Centro Escolar Depto. Gestión de Calidad  
Aprobado por: Centro Escolar Depto. Apoyo al Posgrado

Código: CC-011-010  
Actualización: D  
Emisión: 17/05/18



DICTAMEN DE LIBERACION ACADEMICA PARA INICIAR LOS TRAMITES DEL EXAMEN DE GRADO



Fecha de dictaminación dd/mm/aa: 23/06/20

NOMBRE: JUAN EDUARDO RIOJA ROMO ID 15321
PROGRAMA: MAESTRÍA EN INFORMÁTICA Y TECNOLOGÍAS COMPUTACIONALES LGAC (del posgrado): MODELOS DE GESTIÓN Y EVALUACIÓN DE SISTEMAS Y TI EN ORGANIZACIONES
TIPO DE TRABAJO: ( ) Tesis ( X ) Trabajo práctico
TÍTULO: ANÁLISIS DEL PROCESO DE SUPERVISIÓN DEL PROYECTO "ENCUESTAS EN HOGARES" DEL INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA (INEGI) PARA CONTROL Y TOMA DE DECISIONES
IMPACTO SOCIAL (señalar el impacto logrado): ANALISIS APLICADO CON UTILIDAD Y VALOR PRACTICO PARA EL CONTROL DEL PROCESO Y APOYO EN LA TOMA DE DECISIONES.

INDICAR SI/NO SEGÚN CORRESPONDA:

Elementos para la revisión académica del trabajo de tesis o trabajo práctico:

- SI El trabajo es congruente con las LGAC del programa de posgrado
SI La problemática fue abordada desde un enfoque multidisciplinario
SI Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI El trabajo demuestra más de una aportación original al conocimiento de su área
SI Las aportaciones responden a los problemas prioritarios del país
SI Generó transferencia del conocimiento o tecnológica
SI Cumpe con la ética para la investigación (reporte de la herramienta antiplagio)
El egresado cumple con lo siguiente:
SI Cumple con lo señalado por el Reglamento General de Docencia
SI Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc)
SI Cuenta con los votos aprobatorios del comité tutorial, en caso de los posgrados profesionales si tiene solo tutor podrá liberar solo el tutor
SI Cuenta con la carta de satisfacción del Usuario
SI Coincide con el título y objetivo registrado
SI Tiene congruencia con cuerpos académicos
SI Tiene el CVU del Conacyt actualizado
SI Tiene el artículo aceptado o publicado y cumple con los requisitos institucionales (en caso que proceda)

En caso de Tesis por artículos científicos publicados

- \_\_\_\_\_ Aceptación o Publicación de los artículos según el nivel del programa
\_\_\_\_\_ El estudiante es el primer autor
\_\_\_\_\_ El autor de correspondencia es el Tutor del Núcleo Académico Básico
\_\_\_\_\_ En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación.
\_\_\_\_\_ Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
\_\_\_\_\_ La aceptación o publicación de los artículos en revistas indexadas de alto impacto

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado

Sí X
No

FIRMAS

Elaboró:

\* NOMBRE Y FIRMA DEL CONSEJERO SEGÚN LA LGAC DE ADSCRIPCIÓN:

Mtro. JORGE EDUARDO MACÍAS LUÉVANO

NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO:

Mtro. JORGE EDUARDO MACÍAS LUÉVANO

\* En caso de conflicto de intereses, firmará un revisor miembro del NAB de la LGAC correspondiente distinto al tutor o miembro del comité tutorial asignado por el Decano

Revisó:

NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO:

DRA. EN C. BIO. HAYDÉE MARTÍNEZ RUVALCABA

Autorizó:

NOMBRE Y FIRMA DEL DECANO:

M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ

Nota: procede el trámite para el Depto. de Apoyo al Posgrado

En cumplimiento con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: .... Cuidar la eficiencia terminal del programa de posgrado y el Art. 105F las funciones del Secretario Técnico, llevar el seguimiento de los alumnos.



## **Agradecimientos**

A la Universidad Autónoma de Aguascalientes, por ser la casa de estudios que, una vez más, me ha brindado la oportunidad de adquirir conocimiento y me ha permitido crecer en el ámbito académico.

Al INEGI, por ser la institución que me abrió las puertas en el ámbito profesional y en la cual, he podido comprender y poner en práctica los conocimientos adquiridos durante los últimos años.

A mi comité tutorial, el Mtro. Jorge Eduardo Macías Luévano, la Dra. Lizeth Itziguery Solano Romo y el Dr. Juan Muñoz López, por compartir su conocimiento dentro y fuera del aula, por su motivación y apoyo para la elaboración de este trabajo.

A mis profesores del plan académico, Mtro. Oswaldo Díaz, Dra. Teresa Lucio, Dra. Paola Reyes, Mtro. José de Jesús Palos, Dra. Laura Garza, Dr. José Manuel Mora, Dr. Carlos Arévalo, Dra. María Dolores Torres, Dr. Raúl Valverde, Prof. Esa Kujansuu, Dr. Luis Eduardo Bautista, Mtro. Luis Enrique Arámbula, por compartir sus conocimientos y experiencia.

A mis compañeros y amigos, por los buenos momentos que nos permiten seguir adelante y cumplir nuestros objetivos.

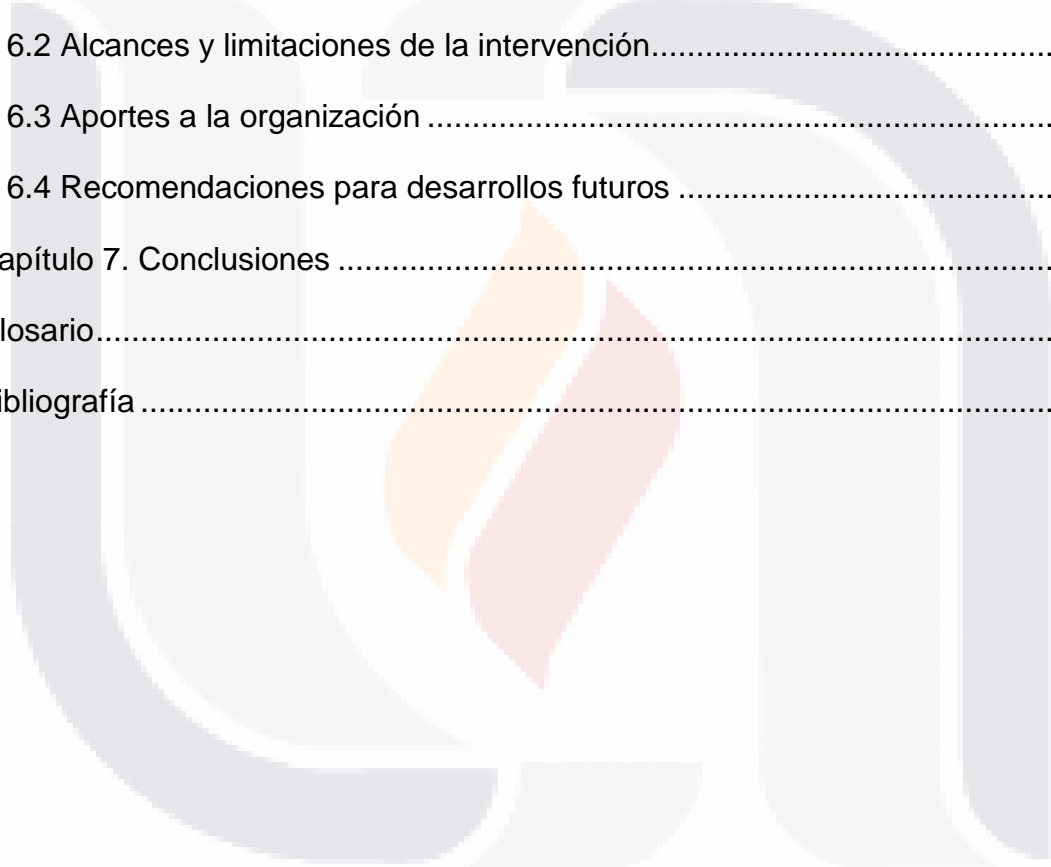
A mi familia y seres queridos con los que he crecido y que continuamos compartiendo experiencias. A mis padres por su cariño, esfuerzo y apoyo incondicional, y a mis hermanos, por estar siempre presentes en mi vida. A mi esposa, por su ejemplo, comprensión y apoyo incondicional, gracias por siempre estar a mi lado. A mis hijos, por la enseñanza y felicidad que nos brindan, son el motor que nos impulsa a ser mejores cada día.

A todos ustedes, ¡Gracias!

## Índice general

|   |    |
|---|----|
| Introducción.....   | 8  |
| Capítulo 1. Planteamiento del problema .....                    | 10 |
| 1.1 Antecedentes.....   | 10 |
| 1.2 Diagnóstico.....  | 12 |
| 1.3 Justificación .....   | 13 |
| 1.4 Sector, población o grupo afectado por la problemática..... | 14 |
| Capítulo 2. Objetivos de la intervención .....                  | 15 |
| Capítulo 3. Marco teórico .....                                 | 16 |
| 3.1 Proceso .....   | 16 |
| 3.2 Control de procesos .....                                   | 18 |
| 3.3 Encuestas en hogares.....                                   | 22 |
| 3.4 Supervisión del proyecto .....                              | 27 |
| 3.5 Toma de decisiones .....                                    | 30 |
| 3.6 Medición del impacto.....                                   | 33 |
| Capítulo 4. Metodología.....                                    | 37 |
| 4.1 Tipo de investigación.....                                  | 39 |
| 4.2 Diseño de investigación.....                                | 40 |
| 4.3 Población objetivo .....                                    | 40 |
| Capítulo 5. Resultados.....                                     | 43 |
| 5.1 Recolección de datos .....                                  | 43 |
| 5.2 Limpieza de datos .....                                     | 47 |
| 5.3 Preparación de datos .....                                  | 48 |
| 5.4 Análisis.....   | 48 |

|   |    |
|---|----|
| 5.4.1 Cobertura.....  | 49 |
| 5.4.2 Desempeño .....   | 55 |
| 5.4.3 Visitas .....   | 66 |
| 5.4.4 Tiempo.....   | 69 |
| Capítulo 6. Evaluación .....                                  | 75 |
| 6.1 Valoración de los objetivos propuestos y alcanzados ..... | 75 |
| 6.2 Alcances y limitaciones de la intervención.....           | 77 |
| 6.3 Aportes a la organización .....                           | 77 |
| 6.4 Recomendaciones para desarrollos futuros .....            | 78 |
| Capítulo 7. Conclusiones .....                                | 79 |
| Glosario.....   | 82 |
| Bibliografía .....  | 84 |



## Índice de tablas

|   |    |
|---|----|
| Tabla 1 Medios de captación utilizados en el proceso de supervisión .....   | 12 |
| Tabla 2. Opciones para la determinación del esquema de muestreo. ....   | 26 |
| Tabla 3. Clasificación en los tipos de decisiones .....   | 31 |
| Tabla 4. Comparación del proceso de toma de decisiones basado en inteligencia artificial y el proceso humano.....         | 32 |
| Tabla 5. Variables de interés por periodo. ....   | 44 |
| Tabla 6. Descripción de tabla de resultados de cobertura. ....  | 44 |
| Tabla 7. Descripción de tabla de resultados de desempeño.....   | 45 |
| Tabla 8. Descripción de tabla de promedio de visitas. ....  | 46 |
| Tabla 9. Descripción de tabla de tiempo promedio de visitas. ....   | 46 |
| Tabla 10. Descripción de tabla de análisis de relación entre entrevistas y reentrevistas.....                             | 47 |
| Tabla 11. Estadísticos de la muestra nacional y porcentaje de reentrevistas captadas. ....                                | 50 |
| Tabla 12. Tabla de distribución de frecuencia de reentrevistas captadas por año.  | 52 |
| Tabla 13. Medidas de distribución de cobertura de reentrevistas y porcentaje de entrevistas logradas. ....                | 53 |
| Tabla 14. Prueba de normalidad para las variables de cobertura de reentrevista y porcentaje de entrevistas logradas. .... | 54 |
| Tabla 15. Correlaciones entre el porcentaje de cobertura de reentrevista y porcentaje de entrevistas logradas. ....       | 55 |
| Tabla 16. Estadísticos descriptivos de reentrevistas logradas en el ámbito nacional. ....                                 | 56 |
| Tabla 17. Tabla de distribución de frecuencia de reentrevistas logradas por año.  | 58 |

Tabla 18. Medidas de distribución de reentrevistas logradas. .... 60

Tabla 19. Prueba de normalidad para la variable porcentaje de reentrevistas logradas por año. .... 60

Tabla 20. Correlación entre reentrevistas y entrevistas logradas por año. .... 61

Tabla 21. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas por año. .... 62

Tabla 22. Coeficientes del modelo de regresión lineal simple..... 63

Tabla 23. Pruebas de normalidad de residuos por año..... 65

Tabla 24. Aplicación de modelo de regresión de entrevistas logradas en función de reentrevistas logradas..... 65

Tabla 25. Prueba de normalidad del promedio de visitas por año. .... 67

Tabla 26. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas y promedio de visitas por año..... 67

Tabla 27. Coeficientes del modelo de regresión lineal simple con dos variables independientes..... 68

Tabla 28. Estadísticos descriptivos de promedio de visitas logradas por año. .... 68

Tabla 29. Prueba t de Student para promedio de entrevistas logradas por año. .. 69

Tabla 30. Prueba de normalidad del tiempo promedio de entrevistas por año. .... 71

Tabla 31. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas y tiempo promedio de visitas por año ..... 72

Tabla 32. Coeficientes del modelo de regresión lineal simple con dos variables independientes..... 72

Tabla 33. Estadísticos descriptivos de tiempo promedio de visitas..... 73

Tabla 34. Prueba t de Student para tiempo promedio de visitas por año..... 73



## Índice de figuras

|  |    |
|--|----|
| Figura 1. Elementos de un proceso. ....  | 17 |
| Figura 2. Representación esquemática de los elementos de un proceso .....                    | 18 |
| Figura 3. Ecuación BPM. ....   | 20 |
| Figura 4. Modo principal de recopilación de datos utilizados en las EPA por región.<br>..... | 24 |
| Figura 5. Ciclo de vida de un proyecto.....  | 28 |
| Figura 6. Dimensiones sugeridas de los indicadores. ....                                     | 35 |
| Figura 7 Ciclo de vida del proyecto .....  | 37 |
| Figura 8. Tasa de variación de muestra nacional para reentrevistas por año.....              | 49 |
| Figura 9. Porcentaje de cobertura nacional de reentrevistas por año. ....                    | 50 |
| Figura 10. Tasa de variación de muestra para reentrevista por oficina y año.....             | 51 |
| Figura 11. Distribución de oficinas con reentrevistas captadas por rango y año. .            | 52 |
| Figura 12. Tasa de variación de reentrevistas logradas por año en el ámbito<br>nacional..... | 56 |
| Figura 13. Porcentaje de reentrevistas por tipo de resultado y año. ....                     | 57 |
| Figura 14. Tasa de variación de reentrevistas logradas por oficina y año. ....               | 58 |
| Figura 15. Distribución de oficinas con reentrevistas logradas por rango y año. ..           | 59 |
| Figura 16. Gráfico de dispersión de residuos.....  | 64 |
| Figura 17. Promedio de visitas por tipo de respuesta y año.....                              | 66 |
| Figura 18. Tiempo promedio de entrevistas por año.....                                       | 70 |

## Resumen

La generación de estadística básica por parte de las Oficinas Nacionales de Estadística se obtiene por medio de proyectos censales, encuestas por muestreo o mediante el aprovechamiento de registros administrativos. En México, el Instituto Nacional de Estadística y Geografía (INEGI) es el organismo responsable de captar y difundir información estadística de México y se rige por una serie de normas y documentos con bases técnicas y metodológicas, los cuales tienen el objetivo de generar información de calidad y confiable.

En INEGI, el proyecto “Encuestas en hogares” implementa varios controles para garantizar que el levantamiento de la encuesta cumpla con la normatividad establecida y permita mantener un nivel de calidad óptimo. Uno de estos controles es el proceso de supervisión que se realiza durante el levantamiento o captación de la encuesta y que tiene como objetivo dar seguimiento y evaluar las actividades realizadas por los entrevistadores.

Como parte del proceso de aseguramiento de la calidad y para dar respuesta a las necesidades detectadas durante los levantamientos previos del proyecto, se han implementado cambios de tipo conceptual y metodológico, por lo que se presenta la necesidad de evaluar el impacto de estos cambios en los resultados obtenidos, no sólo en el proceso de supervisión, sino también en los resultados obtenidos en la encuesta.

El presente trabajo práctico tiene por objetivo ayudar a evaluar el impacto de los cambios realizados al proceso, mediante el análisis de los resultados obtenidos en el primer año de levantamiento con la nueva metodología respecto a los años anteriores, para que en un futuro apoye la mejora continua y la toma de decisiones del área responsable.

## **Abstract**

National Statistics Offices generate basic statistics through census projects, sample surveys or through the use of administrative records. In Mexico, the National Institute of Statistics and Geography (INEGI) is the Institute responsible for capturing and disseminating statistical information from Mexico and is governed by a series of standards and documents with technical and methodological bases, which are intended to generate reliable information.

At INEGI, the “Encuestas en hogares” project implements several controls to ensure that the survey is carried out in accordance with established regulations and allows maintaining an optimum quality level. One of these controls is the monitoring process that is implemented during the survey or collection and which aims to monitor and evaluate the activities of the interviewers.

As part of the quality assurance process and as an answer to the detected needs during the previous surveys of the project, conceptual and methodological changes have been implemented, so there is a requirement to assess the impact of these changes in the results obtained, not only in the monitoring process, but also in the results obtained in the survey.

This practical work aims to be a milestone that assess the impact of the changes made to the process, by analyzing the results obtained in the first year of the survey with the new methodology compared to the previous years, so that in the future it supports the continuous improvement and the decision making process of the responsible area.

## **Introducción**

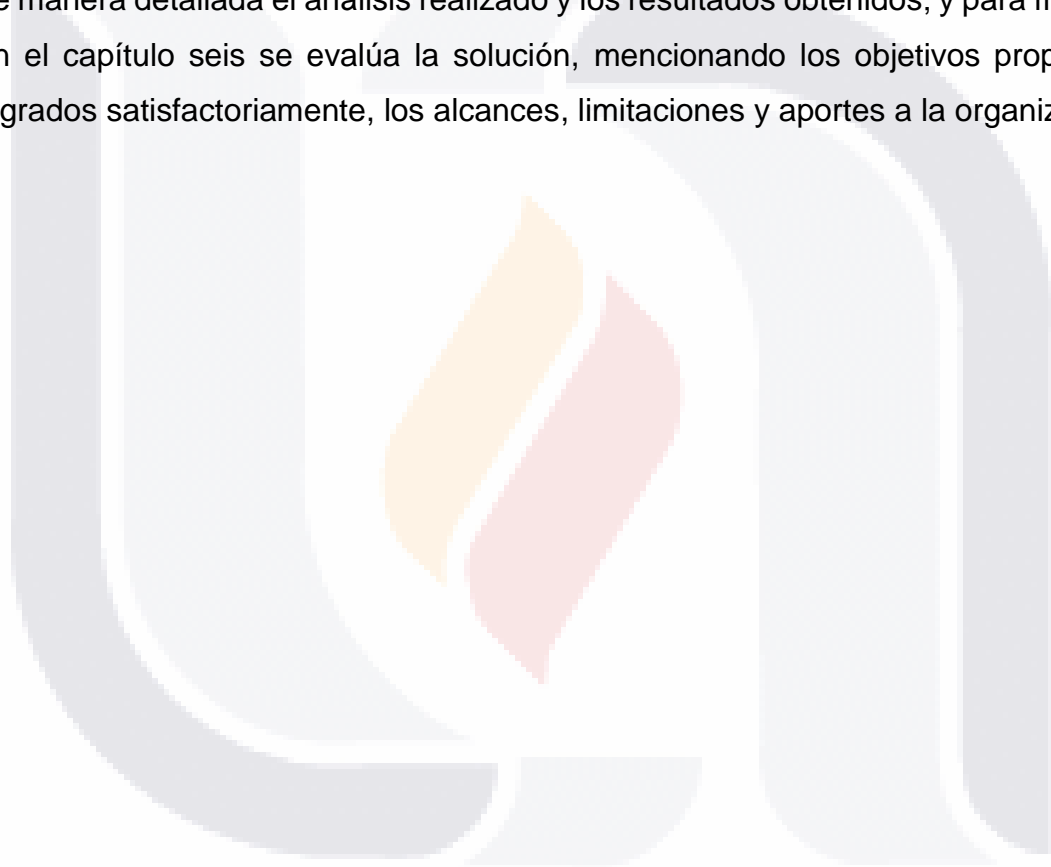
La generación de estadística básica por parte de las Oficinas Nacionales de Estadística, (ONE por las siglas) se obtiene de proyectos censales, de encuestas por muestreo o mediante el aprovechamiento de registros administrativos. En México, el Instituto Nacional de Estadística y Geografía (INEGI) es el organismo responsable de captar y difundir información estadística de México y se rige por una serie de normas y documentos con bases técnicas y metodológicas, los cuales tienen el objetivo de facilitar el cumplimiento de los requisitos de calidad necesarios para la generación de estadística básica (INEGI, 2010).

En INEGI, el proyecto “Encuestas en hogares” implementa varios controles para garantizar que el levantamiento de la encuesta cumpla con la normatividad establecida y permita mantener un nivel de calidad óptimo. Uno de estos controles, es el proceso de supervisión que se realiza durante el levantamiento o captación de la encuesta, el cual tiene como objetivo dar seguimiento y evaluar las actividades realizadas por los entrevistadores.

Como parte del proceso de aseguramiento de la calidad y para dar respuesta a las necesidades detectadas durante los levantamientos previos del proyecto, fue necesario realizar cambios en el proceso de supervisión, como la captación por medio de dispositivos móviles, lo que permitió automatizar la generación de resultados y facilitó la consulta. De manera adicional se han implementado cambios de tipo conceptual y metodológicos en el proceso de supervisión, por lo que se presenta la necesidad de evaluar el impacto de estos cambios en los resultados obtenidos, no sólo en el proceso de supervisión, sino también en los resultados obtenidos en la encuesta.

El presente trabajo práctico tiene por objetivo ayudar a evaluar el impacto, mediante el análisis de los resultados obtenidos en el primer año de levantamiento con la nueva metodología respecto a los años anteriores, para apoyar la mejora continua y la toma de decisiones.

El presente documento se divide en seis capítulos, en los cuáles el primer capítulo describe el panorama actual de la organización donde se lleva a cabo el proyecto, así como la justificación y el sector o área que será beneficiada por el estudio. El capítulo dos describe los objetivos generales y particulares definidos. El capítulo tres presenta el marco teórico que abarca los conceptos clave para el análisis y desarrollo del proyecto. En el capítulo cuatro se describe el diseño de la intervención propuesta para llevar a cabo el análisis del proceso. En el capítulo cinco se describe de manera detallada el análisis realizado y los resultados obtenidos; y para finalizar, en el capítulo seis se evalúa la solución, mencionando los objetivos propuestos logrados satisfactoriamente, los alcances, limitaciones y aportes a la organización.





## Capítulo 1. Planteamiento del problema

### 1.1 Antecedentes

El Instituto Nacional de Estadística y Geografía (INEGI), de acuerdo a lo dispuesto en la Constitución Política de los Estados Unidos Mexicanos, “Es el organismo público con autonomía técnica y de gestión, personalidad jurídica y patrimonio propios, responsable de normar y coordinar el Sistema Nacional de Información Estadística y Geográfica” (SNIEG) (INEGI, 2018, pág. 1).

El INEGI, como filosofía institucional tiene la siguiente misión: “Producir y difundir información, así como normar y coordinar al SNIEG, para suministrar información de calidad, pertinente, veraz y oportuna que coadyuve al desarrollo nacional” (INEGI, 2019). Su visión es: “Ser una institución de vanguardia que satisfaga al usuario de información estadística y geográfica; que contribuya al desarrollo nacional; que asuma el liderazgo del SNIEG y que aproveche el talento de su gente” (INEGI, 2019).

Para poder cumplir sus objetivos, y tomando en cuenta la filosofía institucional, el INEGI ha definido procesos estandarizados que cuidan la calidad de la información y que a su vez promueven la mejora continua y adopción de nuevas tecnologías.

El proyecto “Encuestas en hogares”, desde que inició hasta la fecha ha llevado a cabo un proceso de supervisión a la par de la captación o levantamiento de la encuesta con la finalidad de garantizar la calidad de la información captada. Para lograr esto se cuidan aspectos como la observación a los entrevistadores, donde se verifica que se sigan los protocolos para llevar a cabo la entrevista, que abarcan desde la ubicación correcta de la vivienda hasta la aplicación textual de las preguntas del instrumento de captación.

Otro de los controles realizados es el control de visitas, el cuál registra el número de veces y el tiempo que necesita un entrevistador para lograr obtener la

TESIS TESIS TESIS TESIS TESIS

información sociodemográfica de una vivienda. El tercer control realizado es el de la *reentrevista*, el cuál recaba nuevamente la información captada durante el levantamiento para posteriormente poder generar indicadores de calidad, los cuáles se ponen a disposición de las oficinas estatales para su atención (INEGI, 2018).

Derivado del análisis de los resultados obtenidos en este proceso, y con la finalidad de mejorar la calidad integral de los procesos sustantivos del proyecto, se han realizado diversos cambios, como los metodológicos, que abarcan desde la selección de la muestra de viviendas a supervisar hasta la definición de nuevos controles, como el control de visita, el cual se aplicó por primera vez en el segundo trimestre de 2017.

Los cambios de tipo conceptual se centran en los objetivos de cada control y la información que se obtiene, derivándose en nuevas estrategias para la recolección de la información y nuevos formatos o instrumentos de captación. Por último, durante varios años se han realizado cambios tecnológicos, que abarcan desde la adopción de equipos Asistente Digital Personal (PDA por sus siglas en inglés), captura en línea y, a partir de 2016, el uso de computadoras portátiles (mini laptop).

En la tabla 1 se muestran los distintos medios de captación y los años en los que se utilizaron. También menciona el proceso al que pertenecen, se denominó al proceso original como *tradicional* y al nuevo proceso como *reingeniería*. La transición entre la utilización de los distintos medios de captación fue gradual, para el caso de las PDA, se utilizaron desde 2005 hasta el tercer trimestre de 2014.

A partir del cuarto trimestre y hasta el primer trimestre de 2016, se levantó la información en papel para su captura posterior en un sitio de intranet. En el primer mes del segundo trimestre de 2016, ocho oficinas estatales comenzaron a utilizar equipos mini laptop para la supervisión, mientras que las oficinas restantes continuaron con el levantamiento en papel. En el segundo mes del mismo trimestre se agregaron 15 oficinas para el uso de equipos mini laptop, y a partir del tercer mes

todas las oficinas estatales (38) utilizaron los equipos y dejaron la captura en papel. Dichos equipos se continúan utilizando en 2020.

*Tabla 1 Medios de captación utilizados en el proceso de supervisión*

| <b>Periodo</b> | <b>Medio de captación</b> | <b>Proceso</b> |
|----------------|---------------------------|----------------|
| 2005 - 2014    | PDA                       | Tradicional    |
| 2014 - 2016    | Captura web / Mini laptop | Tradicional    |
| 2017           | Mini laptop               | Tradicional    |
| 2018 - actual  | Mini laptop               | Reingeniería   |

Fuente: Elaboración propia.

Como los cambios conceptuales, metodológicos y tecnológicos ya se realizaron a partir de enero de 2018, ya se cuenta con resultados trimestrales de cobertura y desempeño, pero es necesario analizar los resultados obtenidos, para conocer el impacto que se ha tenido en el mismo proceso, y en los resultados obtenidos en la encuesta principal.

### **1.2 Diagnóstico**

El seguimiento, análisis y generación de resultados que se ha realizado hasta el momento ha sido posible gracias al apoyo de las tecnologías de información. Al momento se generan reportes de manera continua y se presentan a las oficinas estatales por medio de un portal web. El análisis de los resultados obtenidos es constante y el área encargada genera, de manera adicional, informes que retroalimentan para lograr un mejor desempeño de los equipos operativos, los cuales llevan a cabo de manera directa las actividades de captación y supervisión del proyecto.

Tanto la información que se está generando día a día como la histórica se encuentra almacenada en un sistema manejador de base de datos, lo que permite el acceso y generación del análisis solicitado, para evaluar el impacto en los cambios en el proceso de supervisión ya mencionados, derivados de la necesidad de mejora

continua en la calidad del proyecto. Por lo tanto, se pretende desarrollar dicho análisis con el apoyo de herramientas informáticas que cubra esta necesidad y facilite la toma de decisiones en un futuro.

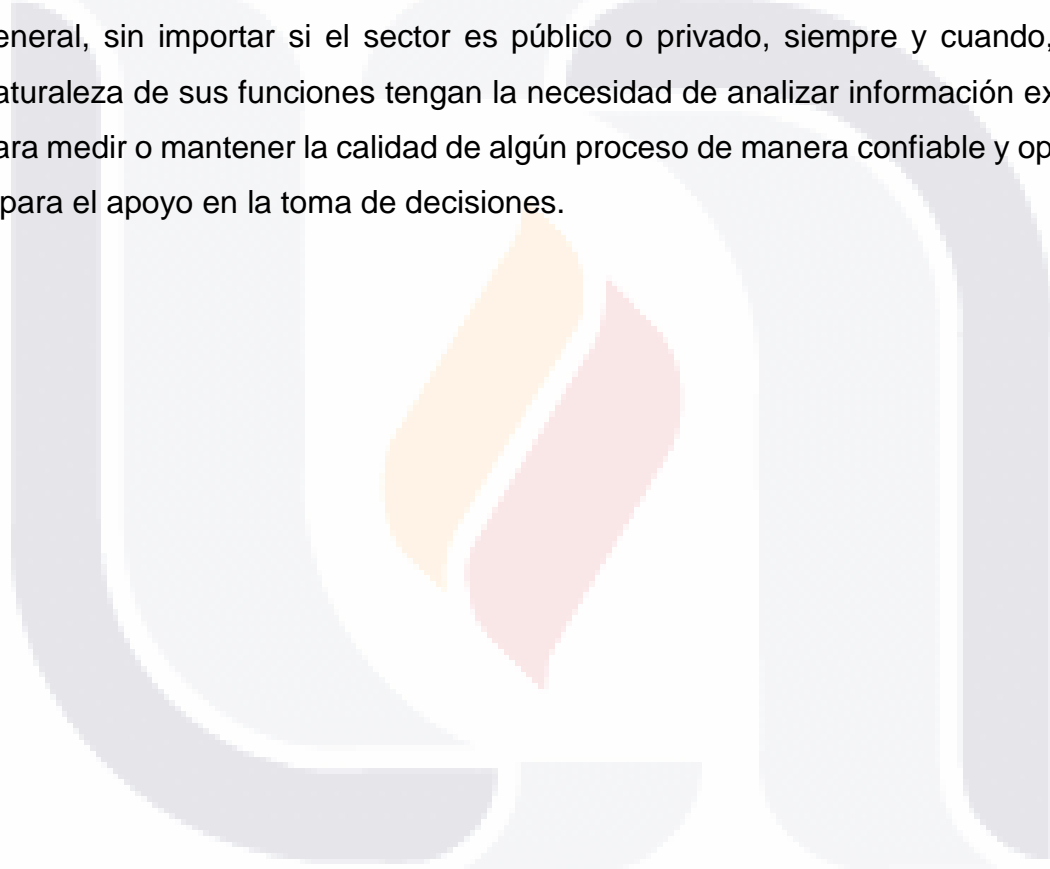
### **1.3 Justificación**

La información del proyecto “Encuestas en hogares” forma parte de la oferta permanente del Instituto Nacional de Estadística y Geografía (INEGI) y tiene el propósito de dar cumplimiento a la Ley del Sistema Nacional de Información Estadística y Geográfica. Para aplicar la mejora continua de la calidad en la ejecución del proyecto, existe un área responsable de la evaluación de los procesos propios del proyecto. Dicha área define los procesos de supervisión en los cuales se detectan áreas de oportunidad y se toman decisiones para estandarizar las actividades del personal operativo encargado tanto de levantar la información como de supervisarla.

El área responsable genera resultados de manera trimestral y a su vez cuenta con resultados históricos del proceso de supervisión, pero se ha presentado la necesidad de llevar a cabo un análisis detallado que, aprovechando las tecnologías de información, permita evaluar el impacto y facilite, en base a los resultados obtenidos, detectar áreas de oportunidad y coadyuvar la toma de decisiones. El proyecto “Encuestas en hogares” en este momento está atravesando un periodo de transición en el cual se esperan cambios de tipo conceptual en el instrumento de captación y el cambio en los medios de captación a dispositivos móviles, por lo que el análisis realizado en el presente trabajo apoyará a la toma de decisiones que permitan llevar un control adecuado en los procesos de control y supervisión durante el periodo de transición y para proyectos posteriores o similares.

#### **1.4 Sector, población o grupo afectado por la problemática**

El área que recibirá el apoyo del trabajo práctico es principalmente el área encargada del proyecto “Encuestas en hogares” del INEGI, ya que contarán con más elementos y herramientas que ayuden a analizar y sustentar las decisiones de los cambios futuros y permitan la mejora continua del seguimiento y de los procesos de supervisión, cumpliendo con los requisitos de calidad ya establecidos en el proyecto. Sin embargo, el conocimiento adquirido lo puede aplicar la sociedad en general, sin importar si el sector es público o privado, siempre y cuando, por la naturaleza de sus funciones tengan la necesidad de analizar información existente para medir o mantener la calidad de algún proceso de manera confiable y oportuna, y para el apoyo en la toma de decisiones.





## Capítulo 2. Objetivos de la intervención

Objetivo general:

- Evaluar el impacto en los cambios de tipo conceptual y metodológico que se han realizado al proceso de supervisión del proyecto “Encuestas en hogares”, mediante el análisis del proceso para detectar áreas de oportunidad y apoyar la mejora continua y toma de decisiones.

Objetivos específicos:

- Identificar las variables que permitan cuantificar y analizar los cambios realizados al proceso de supervisión.
- Realizar el análisis del impacto aplicando métodos estadísticos y aprovechando las tecnologías de información de acuerdo a la normatividad del Instituto.
- Presentar el reporte final del análisis del impacto a la organización.
- Presentar resultados y sugerencias a la organización para apoyar y mejorar la toma de decisiones.

## Capítulo 3. Marco teórico

### 3.1 Proceso

Un proceso se define como “una serie de eventos para producir un resultado” (BPM Forum, 2013). La Real Academia Española (2018) define a un proceso como un “conjunto de las fases sucesivas de un fenómeno natural o de una operación artificial”. De acuerdo a la *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía* se define al proceso como un “conjunto de actividades, recursos humanos, datos e infraestructura relacionadas lógicamente para producir un resultado” (INEGI, 2019, pág. 6).

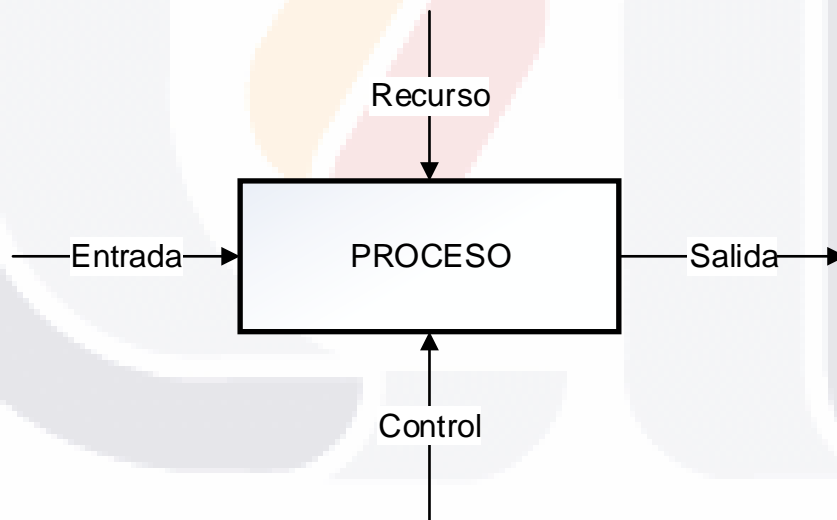
En la *Guía de los fundamentos para la dirección de proyectos (guía del PMBOK)* se define un proceso como “una serie sistemática de actividades dirigidas a producir un resultado final de forma tal que se actuará sobre una o más entradas para crear una o más salidas” (Project Management Institute, 2013, pág. 586).

Bernhard Hitpass (2017) define a un proceso como “Una concatenación lógica de actividades que cumplen un determinado fin, a través del tiempo y lugar, impulsadas por eventos”. También hace énfasis en los principales elementos de un proceso:

- Los eventos son acontecimientos que inician un proceso, es decir, el proceso reacciona ante un evento, algo que sucede.
- Los procesos deben cumplir con un objetivo o deben tener una finalidad específica.
- Las actividades necesarias para llevar a cabo un proceso consumen tiempo y recursos.
- Las actividades en un proceso están relacionadas unas con otras a través de una secuencia lógica u orden, que determinan en su conjunto las condiciones de la organización.

La figura 1 muestra una representación gráfica de los elementos básicos de un proceso, los cuáles se mencionan en las distintas definiciones:

- Entradas: Pueden ser elementos físicos, humanos o técnicos. Son los elementos necesarios para que pueda llevar a cabo un proceso.
- Salidas: Al igual que las entradas, pueden ser tangibles o intangibles, esto debido a que las salidas pueden ser entradas de otros procesos. Es el resultado obtenido, el cuál cumple con el objetivo del proceso.
- Control: Es el elemento que verifica el funcionamiento del proceso y del nivel de satisfacción del usuario, regula cómo, cuándo y si una actividad se ejecuta o no. Pueden ser normas, políticas, calendarios, presupuestos.
- Recursos: También mencionados por algunos autores como mecanismos, son los medios y requisitos necesarios para ejecutar un proceso, por ejemplo: Hardware, Software, recursos humanos e instalaciones.



*Figura 1. Elementos de un proceso.*

Fuente: Elaboración propia basada en (ISO 9001, 2013)

En general existen más de una entrada y más de una salida, frecuentemente una de las salidas puede entrar al mismo proceso, esto es conocido como “Retroalimentación” (ISO 9001, 2013).

Tomando en cuenta las definiciones citadas se puede ver que en el INEGI se hace énfasis a los recursos que intervienen en un proceso, el cual se compone de una serie de eventos, fases o actividades, las cuáles se llevan a cabo en un orden lógico y que tienen como objetivo producir un resultado específico. Las entradas y salidas están implícitas en la definición, pero se describen de manera detallada en el modelo del proceso estadístico y geográfico (MPEG) en el cual se basa el INEGI para la generación de estadísticas.

### 3.2 Control de procesos

La figura 2 muestra una representación de los elementos de un proceso, así como la relación de los controles que pueden existir en cada elemento para dar seguimiento y medir el desempeño, es decir, cada punto de control es específico para cada proceso y varía dependiendo de los riesgos relacionados (International Organization for Standardization (ISO), 2015).



Figura 2. Representación esquemática de los elementos de un proceso

Fuente: (Hitpass, 2017)

La Gestión de Procesos de Negocio (BPM por sus siglas en inglés) es una metodología que orienta la optimización de procesos dentro de una organización por medio de la automatización, integración, optimización y monitoreo continuo (Díaz, 2008, pág. 154).

Los productos y servicios que entrega una organización se derivan de la estrategia del negocio, la cual por lo general se especifica en términos de metas. Los productos y servicios se llevan a cabo por medio de procesos del negocio, los cuáles son monitoreados durante todo el ciclo de vida, mediante la definición de indicadores de rendimiento, calidad y riesgos (BPM Forum, 2013).

Bernhard Hitpass (2017) menciona que las cuatro principales áreas de conocimiento de BPM son las siguientes:

- Negocio: Descripción del modelo del negocio mediante la definición de estrategias, objetivos y metas.
- Proceso: Diseño y mejora de los procesos del negocio.
- Gestión: Ejecutar y controlar la operación.
- Tecnología: Implementar la parte técnica de los procesos y dar soporte a los acuerdos de niveles de servicio (SLA por sus siglas en inglés).

La Figura 3 muestra una ecuación que resume la organización y estructura de la Gestión de Procesos de Negocio (BPM), tomando en cuenta las cuatro principales áreas de conocimiento. Cabe mencionar que cada una de estas áreas representan una especialización en gestión de negocios, en donde se pueden encontrar áreas estratégicas, de finanzas, de operaciones, de calidad y de control de procesos, por mencionar algunas. Todas estas disciplinas se integran en BPM, como un modelo integrado.





*Figura 3. Ecuación BPM.*

Fuente: Elaboración propia basada en (Hitpass, 2017)

La *Norma técnica para la generación de estadística básica* proclamada por la Junta de Gobierno del Instituto define al proceso para la generación de estadística básica como el “Conjunto de procedimientos y actividades para producir estadísticas, a partir de la aplicación de un instrumento de captación a unidades de la población objeto de estudio” (INEGI, 2010, pág. 2). En la norma técnica se definen fases para el proceso de generación de estadística básica y se definen en la misma norma como “Cada una de las series de actividades que se distinguen por su naturaleza técnica específica y los momentos de realización, dado un programa y calendario del proyecto estadístico” (INEGI, 2010, pág. 2).

El INEGI, para llevar a cabo el control de procesos de manera eficiente, se ha regido hasta el año 2018 por el *Proceso estándar para encuestas por muestreo*, el cual define que se debe llevar a cabo un control del operativo, para asegurar que las actividades se realicen en el tiempo y lineamientos establecidos, así como la definición de acciones correctivas para mantener la calidad esperada.

El control operativo se lleva a cabo por las actividades de supervisión tanto del personal como de los resultados y se apoya de las tecnologías de información para obtener dichos resultados. Los aspectos básicos que contempla el proceso son control de cobertura, control de calidad en el llenado de cuestionarios y control de avance (INEGI, 2010, pág. 16).

En 2014, la Asamblea General de la Naciones Unidas adoptó los *Principios Fundamentales de la Estadísticas Oficiales* (Organización de las Naciones Unidas (ONU), 2014), a los cuáles México se adhirió y a finales de 2014 la Junta de Gobierno aprobó los *Principios y Buenas Prácticas para las Actividades Estadísticas y Geográficas del SNIEG* (INEGI, 2015), los cuales están agrupados en tres apartados: entorno institucional, procesos e informantes. Estos tres apartados tienen una relación directa con las áreas de conocimiento de BPM (Negocio, Procesos, Administración (personas) y Tecnología).

A partir de 2017 el INEGI comienza el proceso de adopción del Modelo Genérico del Proceso Estadístico *GSBPM*, por sus siglas en inglés (UNECE, 2013) desarrollado por la Comisión Económica de las Naciones Unidas para Europa (UNECE) y genera el Modelo del Proceso Estadístico y Geográfico *MPEG* (INEGI, 2017)

El MPEG es un modelo conceptual que tiene por objetivo estandarizar el proceso estadístico y facilitar la documentación de los procesos, estandarización de datos y metadatos, reutilización de métodos y componentes, armonización de infraestructuras y aplicaciones de software, y promover la mejora y evaluación continua del proceso de calidad, así como la armonización con el marco normativo aplicable al mismo.

### 3.3 Encuestas en hogares

La Real Academia Española (2018) define a una encuesta como el “conjunto de preguntas tipificadas dirigidas a una muestra representativa de grupos sociales, para averiguar estados de opinión o conocer otras cuestiones que les afectan”.

En el artículo 2 de la *Norma técnica del proceso de producción de información estadística y geográfica para el Instituto Nacional de Estadística y Geografía* se mencionan distintas fuentes de información de las cuáles el INEGI produce información estadística y geográfica: censos, encuestas, registros administrativos, imágenes o la combinación de las mismas (INEGI, 2019, pág. 4).

El INEGI en la *Norma técnica para la generación de estadística básica* define a las encuestas por muestreo como un “método para generar información estadística mediante la captación de datos para un subconjunto de unidades seleccionadas de la población objeto de estudio” (INEGI, 2010, pág. 2).

En resumen, las encuestas son operaciones de recolección de datos que se llevan a cabo para captar información de una muestra del universo de estudio. El INEGI realiza encuestas en hogares y establecimientos cuando se presenta una necesidad de información actualizada y con un nivel de detalle más alto sobre temas específicos. Las encuestas en hogares, las clasifica en regulares, las cuales son parte del programa de trabajo permanente y especiales, las que se llevan a cabo por solicitud de instituciones del sector público (INEGI, 2019).

Las encuestas en hogares fueron la respuesta a la necesidad de estadísticas oficiales más amplias, profundas, menos costosas y con resultados oportunos y relevantes, a comparación de los censos, ya que proveen información con precisión y permiten obtener errores de muestreo (Brown, Citro, House, Marton, & Mackie, 2014).

La agencia del gobierno federal canadiense (Canada, 2011) menciona los siguientes aspectos de las encuestas en hogares:

- **Frecuencia:** Puede ser considerada desde dos perspectivas, la frecuencia en que la información es recolectada y en la que es publicada. La diferencia consiste en que existen varios proyectos en los que la información se capta de manera continua, pero sólo se publica cuando se acumula muestra suficiente y proporciona información confiable.
- **Entrevistador:** Por lo general se necesita que los entrevistadores estén altamente capacitados y cuenten con experiencia en encuestas. Esto es necesario ya que necesitan habilidades específicas como la habilidad de convencer a los informantes para proporcionar la información, manejar cuestionarios complejos o metodologías complejas, como la selección de un informante que cuenta con características específicas.
- **Muestreo:** Para las encuestas en hogares, el interés principal es conocer los atributos de las personas que viven en el hogar. Para esto, contar con un marco de muestreo, un diseño o modalidad de muestreo específico y los métodos de estimación permiten que la información obtenida sea consistente y confiable para las variables de interés de la encuesta.
- **Modos de respuesta:** Se puede obtener respuesta de los informantes por distintos métodos, como lo son la entrevista directa, cara a cara con el informante, utilizando cuestionario en papel y captura posterior. También se puede utilizar un equipo o dispositivo móvil, el cual cuenta con el instrumento de captación o cuestionario digitalizado. Otro método es por medio de entrevistas telefónicas y, por último, la auto numeración, la cual se ha realizado por medio de correo o en los últimos años por acceso a una aplicación web.

Los modos de respuesta dependen del costo y complejidad de la información que se quiere obtener, ya que por ejemplo, la entrevista directa es la más costosa pero la única con la que se puede obtener respuestas más confiables a cuestionarios complejos y por otro lado, la auto numeración es menos costosa pero la tasa de respuesta es más baja que los otros modos de

muestreo. Es común que en varios proyectos se utilice más de un modo de respuesta para lograr obtener el porcentaje de entrevistas deseados.

En conclusión, el uso de entrevistas asistidas por computadora (CAI por las siglas en inglés), tanto para los modos telefónicos o personales ha permitido mejorar la calidad y el control de la productividad para los procesos de entrevistas en hogares. También el uso de los sistemas ha permitido que parte del procesamiento se realice al momento de la entrevista, permitiendo detectar errores y corregirlos. Sin embargo, en los últimos años cada vez es más difícil obtener información vía telefónica por el incremento del uso de telefonía celular, y que cada vez es más común que los hogares no cuenten con una línea telefónica fija.

La figura 4 muestra el total de países y los modos principales de recopilación de datos utilizados en las Encuestas de Población Activa (EPA) por región. Es importante mencionar que en 2018 el 55% de los países que levantan información de población activa utilizan sólo un modo de recolección de información y poco más del 50% utiliza la entrevista personal asistida por computadora (CAPI) como el modo principal, seguido por la entrevista asistida por cuestionario en papel (PAPI) con un 27% y las entrevistas realizadas por internet (CAWI) con un 19%. (Benes & Walsh, 2018).

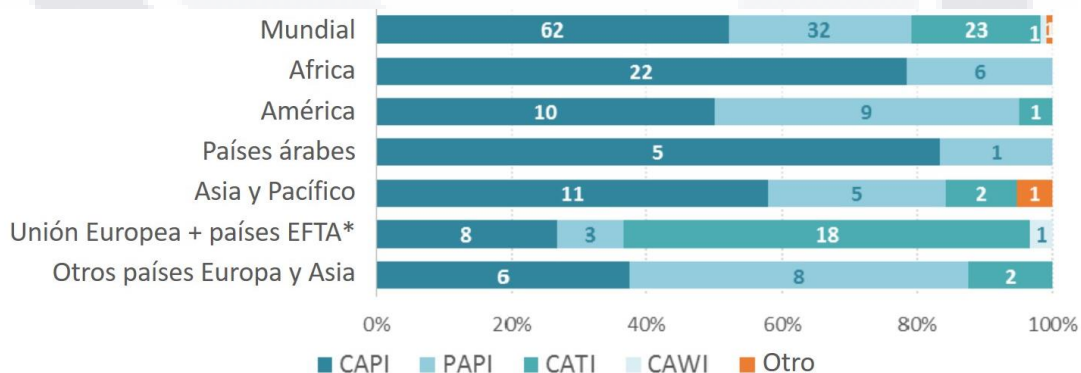


Figura 4. Modo principal de recopilación de datos utilizados en las EPA por región.

Fuente: Elaboración propia basada en (Benes & Walsh, 2018)



Un marco de muestreo se define como “el conjunto de materiales (listas, archivos, mapas, etc.) que permiten identificar a todos los elementos de la población, seleccionar una muestra y localizar sus unidades en campo una vez seleccionadas” (INEGI, 2011).

Una encuesta cuenta con un esquema de muestreo, el cual es una combinación de tipo de muestreo, modalidad de muestreo y número de etapas de selección. El tipo de muestreo puede ser probabilístico, el cual tiene la característica de que cada elemento de la población tiene una probabilidad de ser seleccionado y determinístico, donde se selecciona una muestra bajo criterios subjetivos.

La diferencia entre los dos tipos de muestreo es que con el probabilístico se pueden generalizar resultados a las poblaciones objeto de estudio mientras que en el determinístico no es posible. La modalidad de muestreo es una subclasificación del tipo de muestreo y para determinar la mejor alternativa es necesario analizar los condicionantes y sus respectivas ventajas y desventajas. Por último, hay que considerar el número de etapas de selección.

Los factores que condicionan el esquema de muestreo van desde la amplitud y desglose geográfico hasta la cantidad de variables y el grado de precisión a obtener, y si se cuenta o no con un marco muestral (INEGI, 2011). La tabla 2 muestra las opciones para determinar el esquema de muestreo:

Tabla 2. Opciones para la determinación del esquema de muestreo.

| Tipo de muestreo/Modalidad de muestreo  | Número de etapas de selección                      |  |             |
|---|--|--|-------------|
|   | Unietápico   | Bietápico  | Polietápico |
| <b>Probabilístico</b> <ul style="list-style-type: none"> <li>• Aleatorio simple<sup>1</sup></li> <li>• Sistemático</li> <li>• Estratificado</li> <li>• Por conglomerados</li> </ul><br><b>No probabilístico o determinístico</b> <ul style="list-style-type: none"> <li>• Convencional o accidental</li> <li>• Por cuotas</li> <li>• Cadena o bola de nieve (<i>snow ball</i>)</li> <li>• Intencional o por juicio</li> </ul> | Una sola elección del tipo y modalidad de muestreo | Combinación específica del tipo de muestreo, la modalidad y el número de etapas de selección |             |

<sup>1</sup> Existen dos variantes para esta modalidad diferenciadas por la manera de realizar la selección, con o sin reemplazo.

Fuente: (INEGI, 2011)

Para muchos autores, las encuestas en hogares con representatividad nacional se consideran como una de las innovaciones más importantes en investigaciones de ciencias sociales del último siglo. Éstas permiten medir los errores de estimación, es decir, la diferencia entre el valor estimado y el valor verdadero, pudiendo así cuidar la calidad de la información captada.

La medición del error también se logra comparando la información captada con datos históricos o datos obtenidos de otras fuentes, aunque esto también presenta dificultades como aspectos metodológicos que no permiten que las poblaciones entre las fuentes empaten. Otra dificultad se deriva en la falta de respuesta, existen diversos estudios que demuestran que la no respuesta ha ido aumentando en los últimos años.

Una explicación en la declinación de la obtención de respuestas es el incremento en la urbanización, lo cual disminuye la motivación pública debido al estrés, el aumento del crimen y el aumento de la preocupación por aspectos de privacidad y confidencialidad. Esto se denomina como error de cobertura, el cual aumenta cuando la muestra no representa de manera apropiada a la población objeto de estudio.

Tomando en cuenta estos aspectos se puede decir que la calidad de las encuestas en general va declinando y cada vez es más costoso mantener la calidad de las mismas, por lo que ha ido aumentando el uso de fuentes de información alternativas para la investigación y toma de decisiones políticas, como los registros administrativos (Meyer, Mok, & Sullivan, 2015).

### **3.4 Supervisión del proyecto**

El proceso de supervisión que se realiza en los proyectos de encuestas en hogares es parte fundamental del proceso de captación que se realiza en el Instituto Nacional de Estadística y Geografía (INEGI) y tiene como objetivo evaluar y mantener la calidad de los procesos fundamentales del levantamiento de las encuestas del Instituto.

Como se menciona en el artículo 21 de la *Norma técnica para la generación de estadística básica* proclamada por la Junta de Gobierno del Instituto “El levantamiento de la información debe ser objeto de seguimiento, análisis de avances y atención a desviaciones y contingencias, con apoyo en la supervisión, para aplicar oportunamente acciones correctivas y cumplir con lo programado” (INEGI, 2010, pág. 6).

De acuerdo al *Proceso estándar para encuestas por muestreo* se define al proceso de supervisión como “la observación del desempeño del personal en las distintas funciones, con fines de evaluación y apoyo al mejoramiento de su trabajo” (INEGI, 2010), y se implementa en el proyecto de “Encuestas en hogares” de acuerdo a lo establecido en el artículo 17 numeral III de la *Norma técnica para la generación de estadística básica* “La estrategia para la captación de datos deberá cubrir los siguientes aspectos: Procedimientos para el control de la cobertura y la calidad en el llenado de cuestionarios o formatos de captación, con apoyo en la supervisión” (INEGI, 2010, pág. 5).

De manera adicional, la supervisión es parte de los principios de calidad de los procesos estadísticos y geográficos, en particular el principio de *Implementación adecuada* describe que es necesario “Realizar controles, supervisiones y evaluaciones de las distintas fases del proceso para detectar áreas de mejora” (INEGI, 2015).

La *Guía del PMBOK* describe la naturaleza de la dirección de proyectos por medio de la integración de procesos, su interacción y objetivos. A estos procesos los agrupa en cinco categorías: Procesos de inicio, de planificación, de ejecución, de monitoreo y control y, por último, los procesos de cierre. Es importante hacer énfasis en los procesos de monitoreo y control, los cuáles son “aquellos procesos requeridos para rastrear, revisar y regular el progreso y el desempeño del proyecto, para identificar áreas en las que el plan requiera cambios y para iniciar los cambios correspondientes” (Project Management Institute, 2013).

Los procesos de monitoreo y control interactúan con los otros grupos de procesos y se ejecutan al mismo tiempo que los otros procesos. El mayor beneficio de contar con estos procesos es que permiten supervisar el desempeño, mediante mediciones para realizar análisis de manera regular e identificar desviaciones en un plan de proyecto, controlar cambios en los procesos y recomendar acciones preventivas y correctivas. La figura 5 muestra cómo los procesos de monitoreo y control se encuentran presentes en todo el ciclo de vida de un proyecto:

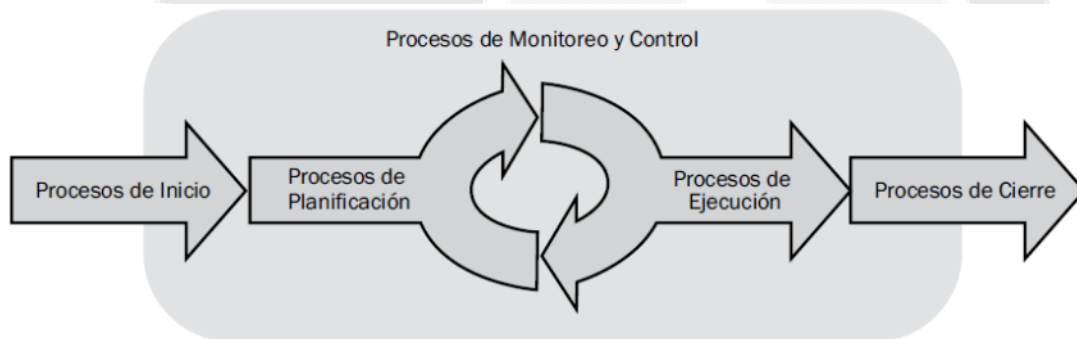


Figura 5. Ciclo de vida de un proyecto.

Fuente: (Project Management Institute, 2013)

El *Documento Metodológico de Supervisión* menciona que, con la finalidad de mantener o elevar la calidad del proyecto “Encuestas en hogares”, se ejecutan procesos de supervisión que permiten analizar y evaluar el desempeño de las oficinas operativas y se generan informes que permiten tomar decisiones y emitir recomendaciones (INEGI, 2018). Para esto se cuenta con distintos instrumentos para cumplir con dicho objetivo:

- **Observación:** Cada entrevistador es observado al menos una vez por semana por un supervisor distinto y se evalúan distintos aspectos de la técnica de entrevista, como si se leyó textualmente la pregunta y si aplicó de manera correcta instrucciones específicas del cuestionario.
- **Reentrevista y control de entrevista:** Permite analizar la correcta identificación de viviendas, condición de residencia de los habitantes del hogar (omisiones o residentes listados por error) y el contenido de los cuestionarios (diferencia en las respuestas). La reentrevista se aplica al ámbito urbano y el control de entrevista al ámbito rural.
- **Control de visitas:** Su propósito es dar seguimiento y llevar un registro detallado de las visitas a las viviendas para localizar a los residentes y registrar la duración de cada una de ellas.

El instrumento que se va a analizar en este estudio es el de la *reentrevista*, ya que es un instrumento “por medio del cual busca establecer un control más preciso sobre la calidad de la información levantada en campo por los entrevistadores” (INEGI, 2018). Entre la entrevista original y la *reentrevista* existen diferencias significativas ya que ésta última tiene objetivos particulares que permiten evaluar la calidad de la encuesta, y no necesariamente obtener nuevamente la encuesta para realizar una comparación que busque una coincidencia exacta en cada una de las respuestas.

Otro aspecto importante es la tecnología utilizada para captar la información del proceso de supervisión. Como se mencionó en los antecedentes del presente

documento, en 2017 se utilizaron equipos mini laptop y a partir de 2018 se utilizaron los mismos equipos, pero con la nueva estrategia, la cual se conoce como la reingeniería del proceso. En la reingeniería, los mayores cambios fueron de tipo conceptual, ya que se utilizaron nuevos instrumentos de captación, y metodológicos, con cambios como el proceso de selección del informante.

### **3.5 Toma de decisiones**

La Real Academia Española (2018) define decisión como la “determinación, resolución que se toma o se da en una cosa dudosa”. Es por esto que se puede definir a la toma de decisiones como el acto de elegir entre dos alternativas o más para lograr un fin específico.

Herbert Simon (1955) simplifica a la toma de decisiones como un proceso racional en el que se realiza una selección de un conjunto de alternativas donde el resultado es satisfactorio. Después se busca para cada alternativa las posibles consecuencias de ese subconjunto seleccionado. Si se selecciona una alternativa derivada de este proceso, entonces se asegura un resultado satisfactorio.

Peñaloza (2010) Afirma que para que exista la toma de decisiones tienen que existir los siguientes elementos:

1. Sujeto o decisor: es la persona que debe tomar una elección.
2. Alternativas o curso de acción a seguir: mínimo deben existir dos, para poder llevar a cabo una elección o decisión.
3. Criterio: base para determinar la alternativa a seguir.

La tabla 3 muestra la clasificación de los tipos de decisiones tomando como criterio la información disponible y muestra el riesgo asociado:



Tabla 3. Clasificación en los tipos de decisiones

| Tipo de decisión             | Información                      | Riesgo |
|------------------------------|----------------------------------|--------|
| Condiciones de certeza       | Clara, exacta y completa         | Bajo   |
| Condiciones de riesgo        | Con probabilidades de ocurrencia | Medio  |
| Condiciones de incertidumbre | Muy escasa o nula                | Algo   |

Fuente: Elaboración propia basada en (Peñaloza, 2010)

La creación de herramientas para la toma de decisiones tiene un fin específico: conducir a la mejora del diseño y operación de algún proyecto, programa o proceso. Esto por lo general se ve reflejado en información (reportes, indicadores) que no sólo permiten medir la calidad, sino que lo hacen de manera oportuna, permitiendo dar seguimiento a la gestión y resultados o metas establecidos (CONEVAL, 2014, pág. 68). Las herramientas de apoyo para la toma de decisiones requieren del análisis de los responsables o expertos en un tema en particular para fundamentar sus decisiones.

Ben-Menahem, Shrestha y von Krogh (2019) realizan una comparación de los procesos de toma de decisiones entre humanos e inteligencia artificial y definen un marco de trabajo con cinco dimensiones:

1. **Espacios de búsqueda:** se refiere a qué tan específicos son los espacios de búsqueda y las funciones objetivos, es decir, el número de alternativas y los criterios para seleccionar la más adecuada.
2. **Interpretación del proceso de toma de decisiones y resultados:** el nivel en el que las decisiones pueden ser interpretadas y explicadas.
3. **Tamaño del conjunto de alternativas:** capacidad para evaluar de manera uniforme un grupo de alternativas.
4. **Velocidad:** el tiempo necesario para llevar a cabo el proceso de toma de decisiones.
5. **Replicación:** Capacidad de replicar el proceso de toma de decisiones y los resultados obtenidos, tomando en cuenta factores que pueden ser internos y externos.

La tabla 4 muestra la comparación entre los procesos de toma de decisiones basados en inteligencia artificial y los realizados por humanos:

*Tabla 4. Comparación del proceso de toma de decisiones basado en inteligencia artificial y el proceso humano*

| <b>Condiciones para la toma de decisiones</b> | <b>Inteligencia artificial</b>  | <b>Humano</b>   |
|---|---|---|
| Espacio de búsqueda para decisiones           | Requiere un espacio de búsqueda de decisiones bien definido con funciones objetivo específicas.                                     | Libertad en la definición del espacio de búsqueda.  |
| Interpretación del proceso y resultados       | La complejidad de las formas funcionales puede hacer que sea difícil interpretar el proceso de toma de decisiones y sus resultados. | Las decisiones se pueden explicar e interpretar, aunque son vulnerables para darles un sentido retrospectivo.   |
| Tamaño del conjunto de alternativas           | Puede implementar un gran conjunto de alternativas.   | Capacidad limitada para evaluar de manera uniforme un gran conjunto de alternativas.  |
| Velocidad                                     | Rápido respecto a las decisiones humanas. No hay diferencia entre velocidad y precisión.  | Lento respecto a la inteligencia artificial. Sí existe diferencia entre la velocidad y la precisión.  |
| Replicación                                   | El proceso y resultados se pueden replicar debido a que depende de procedimientos computacionales estándar.                         | Replicar un proceso de decisión es vulnerable a factores internos como la experiencia, atención, contexto y estado emocional del tomador de decisiones. |

Fuente: Elaboración propia basada en (Ben-Menahem, Shrestha, & von Krogh, 2019)

Tomando en cuenta estas dimensiones se llega a la conclusión de que hasta el momento es muy difícil replicar el comportamiento humano, por lo que es necesario diseñar sistemas híbridos que exploten las ventajas de la inteligencia artificial y la de los tomadores de decisiones para lograr maximizar los resultados esperados (Ben-Menahem, Shrestha, & von Krogh, 2019).

La toma de decisiones es una actividad inherente en el proceso de supervisión del proyecto “Encuestas en hogares”. En base a los resultados y comportamientos observados por las distintas figuras operativas que participan en el proceso, se deben tomar decisiones para cada uno de los ámbitos. Los lineamientos por seguir

se definen por el área responsable del proceso de supervisión desde las Oficinas centrales (sede Aguascalientes) hacia las oficinas estatales (cobertura nacional).

En base al análisis de los resultados generados se deben tomar decisiones que abarcan desde la definición de la metodología hasta procesos operativos como la capacitación al personal, re-instrucciones, seguimiento y generación de nuevos indicadores.

Al momento el área responsable genera indicadores y resultados tabulares que representan el principal insumo para la toma de decisiones, pero se presenta la necesidad de contar con más elementos que, con el apoyo de las tecnologías de información e infraestructura con que cuenta el Instituto, apoyen a la mejora continua de los procesos de control y toma de decisiones.

Otro aspecto adicional es la adopción, por parte del Instituto, de marcos de trabajo y estándares internacionales para la generación de estadística básica, ya que en dichos modelos se propone la definición estandarizada y el uso constante de métricas que permitan evaluar todos los procesos, entre los cuales se encuentra el proceso de supervisión a analizar.

### **3.6 Medición del impacto.**

En el *Glosario de los principales términos sobre evaluación y gestión basada en resultados* se define Impactos (impacts en inglés) como “Efectos de largo plazo positivos y negativos, primarios y secundarios, producidos directa o indirectamente por una intervención para el desarrollo, intencionalmente o no” (CAD, 2010, pág. 24).

En el mismo glosario se define al impacto en el desarrollo institucional como la medida en que una intervención mejora o debilita la capacidad de hacer uso más eficiente, equitativo y sostenible de recursos humanos, financieros y naturales. Esto se puede lograr a través de una mejor definición de procedimientos estables, transparentes, eficaces y previsibles.

TESIS TESIS TESIS TESIS TESIS

Impacto es el cambio o conjunto de cambios duraderos que se producen en un proceso específico mejorando sus indicadores, como resultado de la ejecución de acciones de investigación, innovación y desarrollo que introducen valor agregado a los productos, servicios, procesos y tecnologías. También menciona que el objetivo de medir el impacto es “evaluar el beneficio tangible, la repercusión del resultado y no el resultado en sí” (Quevedo, Chía, & Rodríguez, 2014).

La medición del impacto comprende la evaluación de los métodos y procesos para identificar factores que inciden en los resultados obtenidos. En la *Norma técnica para la generación de estadística básica* se define evaluación como el “Conjunto de procedimientos de seguimiento y control, que permiten la verificación y medición del cumplimiento de los objetivos establecidos en un proyecto estadístico y cada una de sus fases” (INEGI, 2010, pág. 2).

Para el proceso de supervisión del proyecto “Encuestas en hogares” se contemplan técnicas como la confronta de datos, lo cual se define como el proceso de comparación de datos que generalmente se han derivado de diferentes encuestas u otras fuentes, especialmente de aquéllas con diferentes frecuencias, con el fin de asegurar su coherencia y las causas de cualquier diferencia identificada.

La confronta de datos también puede llevarse a cabo entre estadísticas producidas en diferentes países. Estos procesos pueden o no cuantificar el impacto de cualquier diferencia identificada (Statistical Working Group, 2018, pág. 28).

La comparación de datos, entre periodos específicos ya se realiza en los informes generados por el área responsable del proceso, de manera adicional lo que se busca con este trabajo es realizar un análisis que permita identificar y medir el beneficio o repercusión de los cambios realizados al proceso.

El *Manual para el diseño y la construcción de indicadores* (CONEVAL, 2014) menciona cuatro dimensiones de los indicadores para medir el desempeño: eficacia, eficiencia, economía y calidad. Una dimensión es el aspecto del logro del objetivo

que se desea cuantificar. Cada una de las dimensiones tienen las siguientes características:

- Eficacia: Miden el grado de cumplimiento del objetivo establecido.
- Eficiencia: Miden la relación entre el logro y los recursos (económicos, humanos y materiales) utilizados para su cumplimiento.
- Economía: Miden la capacidad para administrar, generar o movilizar de manera adecuada los recursos financieros.
- Calidad: Miden los atributos, capacidades o características de los bienes o servicios que se producen. Permiten monitorear los atributos desde las perspectivas de oportunidad, accesibilidad, percepción de usuarios y precisión.

La figura 6 muestra las dimensiones de acuerdo a lo que se busca medir:

| Cuando se mide: | Se está midiendo: | Los indicadores recomendados:            |
|-----------------|-------------------|--|
| Impacto         | Fin               | • Eficacia                               |
| Resultados      | Propósito         | • Eficacia<br>• Eficiencia               |
| Productos       | Componente        | • Eficacia<br>• Eficiencia<br>• Calidad  |
| Procesos        | Actividades       | • Eficacia<br>• Eficiencia<br>• Economía |
| Insumos         |                   |  |

Figura 6. Dimensiones sugeridas de los indicadores.

Fuente: (CONEVAL, 2014)

Aunque lo ideal sería medir todas las dimensiones, en realidad no es una tarea sencilla ni económica. Los ámbitos de desempeño están representados por lo que

se desea medir, los cuáles los relacionan a un objetivo (lo que se está midiendo) y los indicadores que se recomiendan. Esta recomendación no es limitativa ni exclusiva y los indicadores más adecuados son aquellos que los responsables y operadores perciben que les aportan mayor valor.

En el Modelo del Proceso Estadístico y Geográfico (MPEG) que adoptó INEGI, en la etapa de evaluación menciona que es necesario reunir los insumos necesarios para generar un reporte de evaluación, el cual servirá para la toma de decisiones, acordar un plan de acción e “incluir la consideración de un mecanismo para monitorear el impacto de dichas acciones que, al mismo tiempo, proveen insumos para evaluaciones de iteraciones futuras del proceso” (INEGI, 2017, pág. 24).

Para evaluar el impacto en este trabajo práctico el análisis comienza con las variables de cobertura, desempeño, promedio de visitas y tiempo promedio de visitas. La selección de estas variables está basada en la disponibilidad y bajo el supuesto de que estas variables permitirán evaluar el impacto en los procesos de captación y supervisión de la encuesta. Otro factor de decisión es que estas variables tienen relación directa entre el proceso de supervisión y la encuesta principal.

## Capítulo 4. Metodología

En este capítulo se describe cómo se llevó a cabo el trabajo práctico, las características de la investigación realizada y de la población objeto de estudio, así como las actividades que permitieron el logro de los objetivos específicos y general descritos en el capítulo 2.

Debido a que el objetivo y alcance del presente trabajo práctico es la evaluación del impacto, y no un desarrollo de software específico, se tomó la decisión de utilizar un marco de trabajo de elaboración propia, basándose en la experiencia adquirida en procesamiento de encuestas y en otras metodologías flexibles para análisis de datos como CRISP-DM (Cross Industry Standard Process for Data Mining), la cual, aunque está enfocada a esfuerzos de minería de datos, por ser un modelo flexible e iterativo se puede utilizar como base para cualquier análisis de datos.

En la figura 7 se muestra el diagrama con las etapas que comprenden el ciclo de vida del proyecto, el cuál sirvió como marco de referencia para lograr los objetivos definidos.

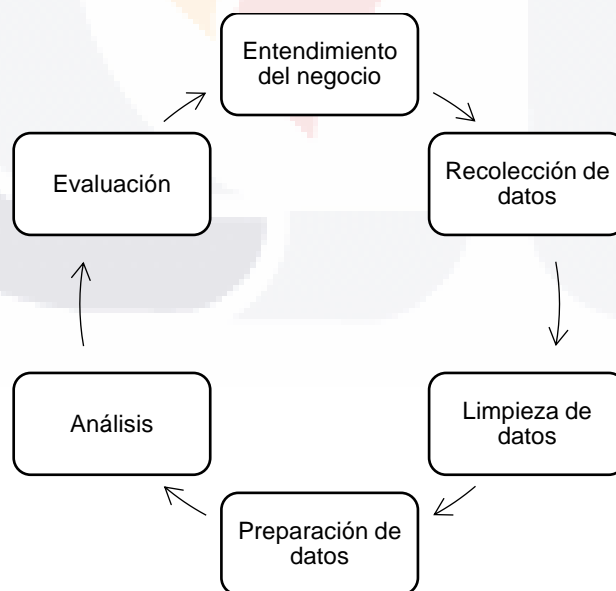


Figura 7 Ciclo de vida del proyecto

Fuente: Elaboración propia, basado en (Smart Vision Europe, 2018)



A continuación, se describe en qué consiste cada uno de los procesos:

1. *Entendimiento del negocio:*

Esta primera etapa comprendió la identificación de la problemática, los antecedentes de la organización, se definieron un diagnóstico y justificación del trabajo práctico y se identificó la población afectada, tomando en cuenta la situación actual de la organización. También se definieron los objetivos del trabajo práctico y se realizó una investigación de temas relacionados con dichos objetivos.

2. *Recolección de datos*

La segunda etapa abarcó la recolección de datos con la finalidad de identificar las distintas fuentes disponibles, y aquellos datos que permitieron realizar un primer análisis para evaluar el impacto. Se identificó su calidad, disponibilidad y relaciones entre variables. Se adecuaron los datos y se describen para facilitar las actividades de los siguientes procesos.

3. *Limpieza de datos*

Durante la integración de datos se identificaron aquellos en los que se presentó alguna inconsistencia, por lo que en esta etapa se descartó que fuera un error con un acercamiento con el área responsable que generó los datos. En los casos que se verificó alguna inconsistencia se tomaron decisiones para que no se afectara la distribución de la información. Se verificaron aspectos como vacíos de información, valores fuera de rango y valores atípicos.

4. *Preparación de datos*

En esta etapa se transformaron los datos para poder generar los primeros análisis estadísticos, se aplicaron criterios de selección, de agrupamiento y se generaron variables adicionales que permitieron observar el comportamiento de los datos en periodos específicos.

## 5. *Análisis*

En esta etapa se generaron y analizaron los resultados para identificar aquellos que permitieron el logro de los objetivos. En esta etapa también se regresó a las etapas de integración, limpieza y preparación en un ciclo iterativo, de acuerdo a los resultados obtenidos.

## 6. *Evaluación*

Con los resultados obtenidos, se realizó una evaluación de los objetivos propuestos y alcanzados, se hizo énfasis en los beneficios obtenidos y se realizaron conclusiones tomando en cuenta los alcances y limitaciones del proyecto.

### **4.1 Tipo de investigación**

Para este trabajo se eligió una investigación de tipo no experimental y evaluativa. Se tomó esta decisión ya que los datos utilizados para el análisis ya se encontraban disponibles para periodos específicos y el objetivo del análisis fue conocer la medida del impacto entre los periodos seleccionados.

Como se mencionó en el marco teórico, se seleccionaron las variables de cobertura y desempeño, así como el promedio de visitas y el tiempo promedio de visitas a la vivienda en periodos anuales. En el proceso de análisis se buscó relacionar los resultados de dichas variables.

De igual manera existen otras variables que se pueden comparar, ya que el objetivo de la *reentrevista* es medir la calidad de la información captada por los entrevistadores. Cabe mencionar que este tipo de análisis comparativo ya lo realiza el área responsable y el objetivo de este trabajo no es evaluar o replicar los resultados ya generados, sino coadyuvar a medir el impacto, a partir de los datos disponibles.

## 4.2 Diseño de investigación

El diseño de la investigación seleccionado en este trabajo fue el longitudinal, se utilizó este diseño ya que los datos recolectados son anualizados para ser comparados. Las cuatro variables seleccionadas para el análisis son cuantitativas y representan el total de viviendas a las que se les aplicó el control de *reentrevista*, las cuáles a su vez, son seleccionadas aleatoriamente del total de viviendas con entrevista completa de la encuesta principal.

- Unidad de análisis primaria: Viviendas seleccionadas para aplicar el control de *reentrevista*.
- Tipo de muestra: Probabilística, muestreo aleatorio sistemático. Representa al 10% de las viviendas logradas en el área urbana, se genera por semana y una vez seleccionada una vivienda para *reentrevista*, no se vuelve a seleccionar para el siguiente periodo (trimestral).

## 4.3 Población objetivo

Para la evaluación se contemplaron todas las viviendas seleccionadas para la aplicación de la *reentrevista*, las cuales sólo representan al área urbana. Existe otro control de supervisión para el área complementaria y rural, el *control de entrevista*, que es el equivalente a la *reentrevista*. Este control no está dentro del alcance de este trabajo ya que, debido a sus características, no comparte el mismo tipo de selección de muestra que la *reentrevista* y tiene otro tipo de temporalidad (mensual).

## 4.4 Técnicas estadísticas

A continuación, se describen las técnicas estadísticas utilizadas para el análisis del presente trabajo práctico.

- Ancho de clase (C): Se obtiene de la división del rango entre el número de clases deseado.

- Amplitud o rango: Es la diferencia entre el valor máximo y mínimo de una variable.
- Asimetría: Medida de distribución que permite identificar si los datos se distribuyen de forma uniforme alrededor del punto central o media aritmética.
- Curtosis: Medida de distribución que determina el grado de concentración de los valores en la región central de la distribución.
- Coeficiente de correlación de Spearman: Medida de correlación ( asociación o interdependencia) entre dos variables aleatorias (continuas o discretas). El coeficiente oscila entre 0 y 1.
- Fórmula de Sturges (K): Sugiere el número de clases con las que se pueden agrupar datos.

$$K = 1 + 3.322 \log N$$

- Media: es el valor promedio de un conjunto de datos numéricos, se calcula mediante la suma del conjunto de valores dividida entre el número total de valores.
- Mínimo: Menor valor de un conjunto de datos.
- Máximo: Mayor valor de un conjunto de datos.
- Pruebas no paramétricas: Tipo de pruebas que no presuponen una distribución específica para la población. Se pueden utilizar con muestras pequeñas asimétricas o con varios valores atípicos.
- Porcentaje: Expresa un número como partes de cada cien. Es el cociente entre dos variables con una misma unidad de medida en el mismo periodo. Permite conocer el grado de avance o cobertura de alguna variable.  
Se calcula de la manera siguiente:

$$\frac{\text{Numerador}}{\text{Denominador}} * 100$$

- Regresión lineal: Existen dos tipos de modelos de regresión de acuerdo a los objetivos buscados:

- Modelos explicativos: El objetivo es evaluar cómo afecta el cambio en características específicas (variables independientes) sobre otra característica determinada (variable dependiente).
- Modelos predictivos: El objetivo es estimar el valor de una variable (dependiente) en función de los valores que pueden tomar en conjunto una o más variables (independientes).
- Tasa de variación: Expresa un cambio relativo en el tiempo. Es el cociente de una misma variable en distintos periodos. Permite medir un incremento o decremento de alguna variable en dos periodos. Se calcula de manera siguiente:

$$\left[ \left[ \frac{\text{Numerador (inf. reciente)}}{\text{Denominador (inf. anterior)}} \right] - 1 \right] * 100$$

#### 4.5 Recursos

Los recursos con los que realizó el análisis son los siguientes:

- Información proporcionada por el área responsable, informes, acceso a sitio con acceso restringido y acceso a bases de datos.
- Software gratuito para consulta de bases de datos.
- Paquete estadístico IBM SPSS Statistics 25.
- Microsoft Office 365 ProPlus
- Hardware propiedad del desarrollador.

## Capítulo 5. Resultados

### 5.1 Recolección de datos

Una vez realizadas las actividades que comprendieron el entendimiento del negocio, se llevó a cabo la identificación y recolección de información que permitió analizar el impacto a partir de los resultados obtenidos. Al momento existen informes y tabulados que se pueden consultar mediante un sitio de intranet donde se publican los resultados y se establece contacto con las oficinas estatales. El área responsable proporcionó el acceso a la información generada, por lo que una vez obtenida dicha información, se realizaron procesos de transformación para facilitar el análisis.

La información de las variables seleccionadas para el análisis se obtuvo de los “Informes de supervisión” y de los resultados generados periódicamente para evaluar el proceso de supervisión. Es aquí donde se detectó uno de los cambios realizados a partir de 2018, los informes de supervisión se realizan de manera trimestral, mientras que en 2016 y 2017 los informes fueron anuales, por lo que se tomó la decisión de realizar el análisis con datos anuales.

La información de 2016 y 2017 se obtuvo de los informes publicados, por lo que la información se extrajo de archivos en formato pdf. A partir de 2018 la información se obtuvo del sitio donde se publican los resultados en formato csv para cada trimestre. La información disponible de cada variable y sus periodos se integraron en hojas de cálculo.

La generación de información anualizada implicó integrar los datos de cada trimestre de 2018 y 2019 de las variables de interés, por lo que en esta etapa se descargaron e identificaron los datos de cada trimestre de los años mencionados.

Sólo la información de visitas de 2017 se tomó también del sitio en formato csv y no del informe, ya que en el informe no se encuentra desagregada la información de la

misma manera que la información disponible de 2018 y 2019. Cabe mencionar que la información de visitas de 2017 es la del cuarto trimestre, por lo que para este periodo no se realizó un proceso adicional para calcular la información anual.

La tabla 5 muestra las variables, periodos anuales y proceso de supervisión de donde se obtuvo la información:

*Tabla 5. Variables de interés por periodo.*

| <b>Variable</b>     | <b>Periodo</b> | <b>Proceso</b>     |
|---------------------|----------------|--------------------|
| Cobertura           | 2016 - 2019    | Reentrevista       |
| Desempeño           | 2016 - 2019    | Reentrevista       |
| Promedio de visitas | 2017 - 2019    | Control de visitas |
| Tiempo promedio     | 2017 - 2019    | Control de visitas |

Fuente: Elaboración propia.

Como resultado, se generaron tablas para cada variable, en las cuáles se integró la información de acuerdo a la siguiente descripción:

*Tabla 6. Descripción de tabla de resultados de cobertura.*

| <b>Variable</b> | <b>Descripción</b>                        | <b>Tipo</b> | <b>Longitud</b> |
|-----------------|---|-------------|-----------------|
| per             | Periodo anual                             | Numérico    | 4               |
| id_oficina      | Identificador operativo de oficina        | Numérico    | 2               |
| desc_oficina    | Nombre de la ciudad                       | Texto       | 45              |
| muestra         | Muestra seleccionada para reentrevista    | Numérico    | 5               |
| captado         | Total de reentrevistas realizadas         | Numérico    | 5               |
| captado_p       | Porcentaje de reentrevistas realizadas    | Decimal     | 2,1             |
| faltante        | Total de reentrevistas no realizadas      | Numérico    | 4               |
| faltante_p      | Porcentaje de reentrevistas no realizadas | Decimal     | 2,1             |

Fuente: Elaboración propia.



Tabla 7. Descripción de tabla de resultados de desempeño.

| <b>Variable</b> | <b>Descripción</b>  | <b>Tipo</b> | <b>Longitud</b> |
|-----------------|---|-------------|-----------------|
| per             | Periodo anual   | Numérico    | 5               |
| id_oficina      | Identificador operativo de oficina                          | Numérico    | 2               |
| desc_oficina    | Nombre de la ciudad   | Texto       | 45              |
| Total           | Total de reentrevistas                                      | Numérico    | 5               |
| res_00          | Reentrevistas logradas                                      | Numérico    | 5               |
| res_00_p        | Porcentaje de reentrevistas logradas                        | Decimal     | 2,1             |
| res_01          | Nadie en el momento de las visitas                          | Numérico    | 4               |
| res_01_p        | Porcentaje Nadie en el momento de las visitas               | Decimal     | 2,1             |
| res_02          | Ausente temporalmente                                       | Numérico    | 3               |
| res_02_p        | Porcentaje Ausente temporalmente                            | Decimal     | 2,1             |
| res_03          | Se negó a dar información                                   | Numérico    | 4               |
| res_03_p        | Porcentaje se negó a dar información                        | Decimal     | 2,1             |
| res_04          | Informante inadecuado                                       | Numérico    | 4               |
| res_04_p        | Porcentaje informante inadecuado                            | Decimal     | 2,1             |
| res_05          | Otro motivo   | Numérico    | 3               |
| res_05_p        | Porcentaje otro motivo                                      | Decimal     | 2,1             |
| res_BC          | Vivienda deshabitada y vivienda fuera de muestra            | Numérico    | 3               |
| res_BC_p        | Porcentaje vivienda deshabitada y vivienda fuera de muestra | Decimal     | 2,1             |
| nr              | Total de no respuesta                                       | Numérico    | 4               |
| nr_p            | Porcentaje total de no respuesta                            | Decimal     | 2,1             |

Fuente: Elaboración propia.

Tabla 8. Descripción de tabla de promedio de visitas.

| Variable     | Descripción                                 | Tipo     | Longitud |
|--------------|---|----------|----------|
| per          | Periodo anual                               | Numérico | 5        |
| id_oficina   | Identificador operativo de oficina          | Numérico | 2        |
| desc_oficina | Nombre de la ciudad                         | Texto    | 45       |
| vis_00       | Promedio de visitas de entrevistas logradas | Decimal  | 1,1      |
| vis_nr       | Promedio de visitas de no respuesta         | Decimal  | 1,1      |
| vis_01       | Nadie en el momento de las visitas          | Decimal  | 1,1      |
| vis_02       | Ausente temporalmente                       | Decimal  | 1,1      |
| vis_03       | Se negó a dar información                   | Decimal  | 1,1      |
| vis_04       | Informante inadecuado                       | Decimal  | 1,1      |
| vis_05       | Otro motivo                                 | Decimal  | 1,1      |
| vis_14       | El hogar se mudó                            | Decimal  | 1,1      |
| vis_15       | Entrevista suspendida                       | Decimal  | 1,1      |
| vis_B        | Vivienda deshabitada                        | Decimal  | 1,1      |
| vis_C        | Vivienda fuera de muestra                   | Decimal  | 1,1      |

Fuente: Elaboración propia.

Tabla 9. Descripción de tabla de tiempo promedio de visitas.

| Variable     | Descripción                                | Tipo     | Longitud |
|--------------|--|----------|----------|
| per          | Periodo anual                              | Numérico | 4        |
| id_oficina   | Identificador operativo de oficina         | Numérico | 2        |
| desc_oficina | Nombre de la ciudad                        | Texto    | 45       |
| promedio     | Tiempo promedio de cada visita realizada   | Decimal  | 2,1      |
| logradas     | Tiempo promedio de entrevistas logradas    | Decimal  | 2,1      |
| no_logradas  | Tiempo promedio de entrevistas no logradas | Decimal  | 2,1      |

Fuente: Elaboración propia.

En la segunda iteración se integraron las siguientes variables:

*Tabla 10. Descripción de tabla de análisis de relación entre entrevistas y reentrevistas.*

| <b>Variable</b> | <b>Descripción</b>                          | <b>Tipo</b> | <b>Longitud</b> |
|-----------------|---|-------------|-----------------|
| per             | Periodo anual                               | Numérico    | 4               |
| id_oficina      | Identificador operativo de oficina          | Numérico    | 2               |
| desc_oficina    | Nombre de la ciudad                         | Texto       | 45              |
| cob_p           | Porcentaje de cobertura captado             | Decimal     | 2,1             |
| ent_p           | Porcentaje de entrevistas logradas          | Decimal     | 2,1             |
| res_00_p        | Porcentaje de reentrevistas logradas        | Decimal     | 2,1             |
| vis_00          | Promedio de visitas de entrevistas logradas | Decimal     | 1,1             |
| promedio        | Tiempo promedio de cada visita realizada    | Decimal     | 2,1             |
| logradas        | Tiempo promedio de entrevistas logradas     | Decimal     | 2,1             |

Fuente: Elaboración propia.

## 5.2 Limpieza de datos

Durante el proceso de integración y transformación de datos se detectaron algunos valores atípicos, los cuáles se reportaron al área responsable. Derivado de este acercamiento, el área responsable del proceso sugirió utilizar la información que se genera en el sitio (tabulados), en lugar de extraer la información de los informes en pdf.

Otro aspecto que se detectó durante la limpieza e integración de datos es que en los periodos analizados ha cambiado el número de oficinas a supervisar, por lo que en distintos años o trimestres se puede encontrar o no alguna oficina. Para efectos del análisis nacional se tomaron en cuenta y en la obtención de resultados se valoró si sería viable generar un análisis por oficina y en su caso, cómo se tratarían estos casos especiales.

En la primera iteración de este proceso se cambiaron los valores N/A (no aplicables) por valores en blanco, esto para facilitar el proceso de preparación de datos.

En cada iteración se omitieron oficinas de 2019 que no tuvieron información en los años anteriores de las variables seleccionadas.

### **5.3 Preparación de datos**

Para facilitar el análisis de la información, en la primera iteración se generaron los datos anuales de 2018 y 2019 mediante el cálculo de un promedio simple para obtener los datos porcentuales. El cálculo se realizó en una hoja de cálculo para cada trimestre y para cada variable objeto de estudio. El resultado final se redondeó a un decimal, para mantener el mismo formato de la información de 2016 y 2017.

Para las oficinas en las que no hubo información en algún trimestre, el promedio se realizó sólo de la información disponible. En estos casos la información mínima fue de tres trimestres por oficina.

Para finalizar, se importó la información de cada variable en el paquete estadístico para realizar los primeros análisis.

### **5.4 Análisis**

Para efectos de este estudio se optó por realizar dos ejercicios, uno tomando los promedios nacionales ya publicados y otro, realizando los cálculos a partir de los datos por oficinas operativas, para obtener los estadísticos descriptivos que permitieron analizar la situación de cada variable a nivel nacional.

En este proceso se presentaron algunas diferencias ya que en los datos publicados sólo se encontraron datos con un decimal, y para generar los resultados se tomaron en cuenta todos los decimales generados por las operaciones aritméticas y al final se redondearon a un decimal. Dado que las diferencias encontradas no fueron significativas ( $\pm 0.1$ ) se optó por utilizar los datos publicados por el área responsable.

Para el análisis por oficina, se crearon tablas de distribución de frecuencia por año, pero tomando el límite inferior y superior de 2016 a 2019, para poder comparar los datos de la variable de interés.

Los análisis tuvieron un contexto geográfico, donde primero se abordaron datos nacionales y después por oficinas, y el contexto temporal, que fue para datos anuales.

### 5.4.1 Cobertura

Los resultados de cobertura muestran el avance respecto a la muestra asignada, y representan un aspecto de logro del objetivo, como indicadores de eficacia.

La muestra de viviendas seleccionadas para reentrevista de los periodos 2016 a 2019 tuvo una media de 23 669 viviendas. Cabe mencionar que la muestra varía de acuerdo a las entrevistas logradas en la encuesta principal y con la integración de nuevas oficinas operativas.

La figura 8 muestra la tasa de variación de la muestra nacional seleccionada para reentrevista respecto al año anterior a partir de 2017. La tasa de variación de 2017 respecto a 2016 muestra una variación negativa del 5.2%, en 2018 la diferencia es menor a un punto porcentual y en 2019 la tasa varía positivamente 6.9%.

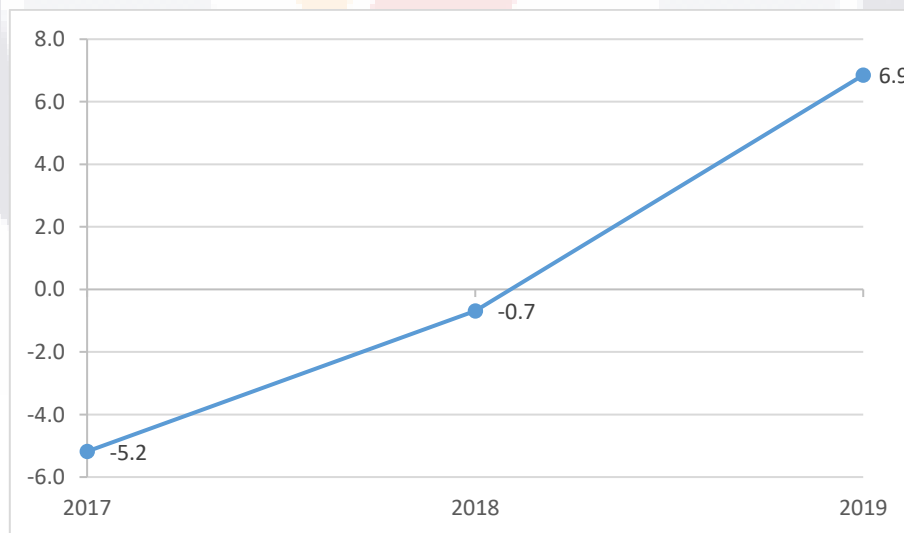


Figura 8. Tasa de variación de muestra nacional para reentrevistas por año.

Fuente: Elaboración propia.

Cada oficina operativa tiene una muestra distinta y tiene relación con su población, por lo que presentan una varianza estadísticamente significativa. La tabla 11 muestra los estadísticos mínimo, máximo, media y desviación típica de la muestra y las reentrevistas captadas de cada año. En los resultados, se puede apreciar que la media de reentrevistas captadas presentó una tendencia positiva, con un aumento respecto al año anterior de 15.9% en 2017, 9.5% en 2018 y 0.9% en 2019.

Tabla 11. Estadísticos de la muestra nacional y porcentaje de reentrevistas captadas.

| Año  | Estadísticos descriptivos |        |        |       |              |                    |        |       |              |
|------|---------------------------|--------|--------|-------|--------------|--------------------|--------|-------|--------------|
|      | Muestra                   |        |        |       |              | Porcentaje captado |        |       |              |
|      | N                         | Mínimo | Máximo | Media | Desv. típica | Mínimo             | Máximo | Media | Desv. típica |
| 2016 | 38                        | 39     | 1574   | 639.5 | 268.3        | 13.5               | 84.7   | 62.7  | 15.4         |
| 2017 | 38                        | 51     | 1465   | 606.3 | 239.6        | 26.2               | 94.1   | 78.6  | 18.5         |
| 2018 | 39                        | 43     | 1356   | 586.7 | 216.9        | 20.2               | 99.5   | 88.1  | 16.2         |
| 2019 | 41                        | 41     | 1371   | 596.4 | 199.2        | 0                  | 99.7   | 89.0  | 20.0         |

Fuente: Elaboración propia.

La figura 9 muestra una tendencia positiva en las viviendas visitadas a nivel nacional en el periodo de 2016 a 2019. Es importante mencionar que, en el periodo de 2016, año en que se comenzó a utilizar los equipos mini laptop, a 2017, hubo un aumento de cobertura del 17.5%. En 2018 se aplicaron los cambios de la reingeniería y también se presentó una diferencia positiva del 9.4% respecto a 2017. En 2019 continúa la tendencia positiva, aunque en menor grado, con un aumento del 1.7%.

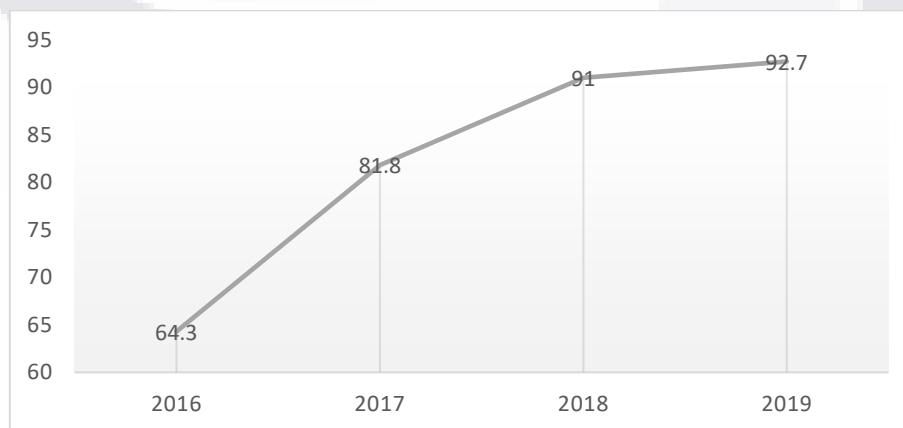


Figura 9. Porcentaje de cobertura nacional de reentrevistas por año.

Fuente: Elaboración propia.

La figura 10 presenta la tasa de variación de la muestra por oficina, lo cual refleja el comportamiento nacional de la figura 8. En 2017 cinco oficinas presentaron un aumento porcentual significativo en la muestra seleccionada, mientras que el resto presentó una diferencia negativa. En 2018 la variación negativa se compensó por el aumento significativo (a nivel porcentaje) de muestra de tres oficinas. En 2019 cuatro oficinas también presentaron un aumento significativo, dos de ellas con una variación mayor al 200% que se muestra en la figura. Dos oficinas que se agregaron en 2019 no se incluyeron en el análisis.

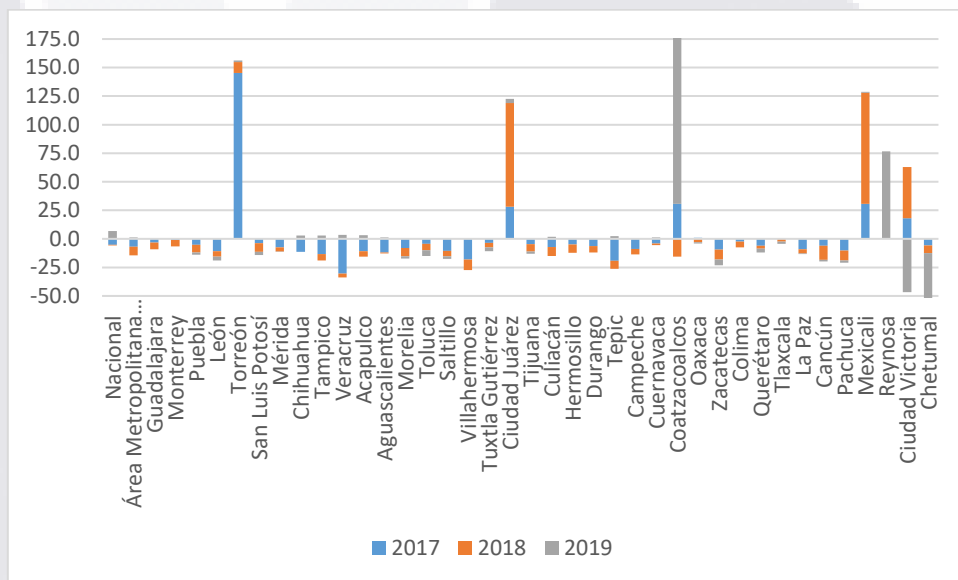


Figura 10. Tasa de variación de muestra para reentrevista por oficina y año.

Fuente: Elaboración propia.

Para facilitar el análisis del porcentaje de cobertura por oficina se creó una tabla de distribución de frecuencia. La tabla 12 muestra las clases (límites inferior y superior (LI,LS), frecuencia de oficinas (F) y frecuencia relativa (FR) de reentrevistas por año. En las frecuencias se puede apreciar la tendencia de aumento del total de oficinas hacia la última clase.



Tabla 12. Tabla de distribución de frecuencia de reentrevistas captadas por año.

| LI    | LS   | 2016 |      | 2017 |      | 2018 |      | 2019 |      |
|-------|------|------|------|------|------|------|------|------|------|
|       |      | F    | FR   | F    | FR   | F    | FR   | F    | FR   |
| 13.5  | 27.9 | 2    | 0.05 | 1    | 0.03 | 1    | 0.03 | 1    | 0.02 |
| 27.9  | 42.3 | 1    | 0.03 | 3    | 0.08 | 1    | 0.03 | 2    | 0.05 |
| 42.3  | 56.7 | 6    | 0.16 | 0    | 0.00 | 0    | 0.00 | 0    | 0.00 |
| 56.7  | 71.1 | 16   | 0.42 | 3    | 0.08 | 1    | 0.03 | 0    | 0.00 |
| 71.1  | 85.5 | 13   | 0.34 | 12   | 0.32 | 6    | 0.15 | 4    | 0.10 |
| 85.5  | 99.9 | 0    | 0.00 | 19   | 0.50 | 30   | 0.77 | 34   | 0.83 |
| Total |      | 38   | 1    | 38   | 1    | 39   | 1    | 41   | 1    |

Fuente: Elaboración propia.

La figura 11 muestra el comportamiento del porcentaje de cobertura obtenido por las oficinas operativas en el periodo de 2017 a 2019. Los intervalos de clase están representados como la variable independiente (eje x) y como variable dependiente el porcentaje de oficinas que se encuentran en dicho intervalo (eje y). A partir de 2017 se muestra que el 50% de las oficinas tuvieron una cobertura mayor al 85%. En 2018 el 77%, y para 2019, el porcentaje de oficinas que se encuentra en este rango es del 83%.

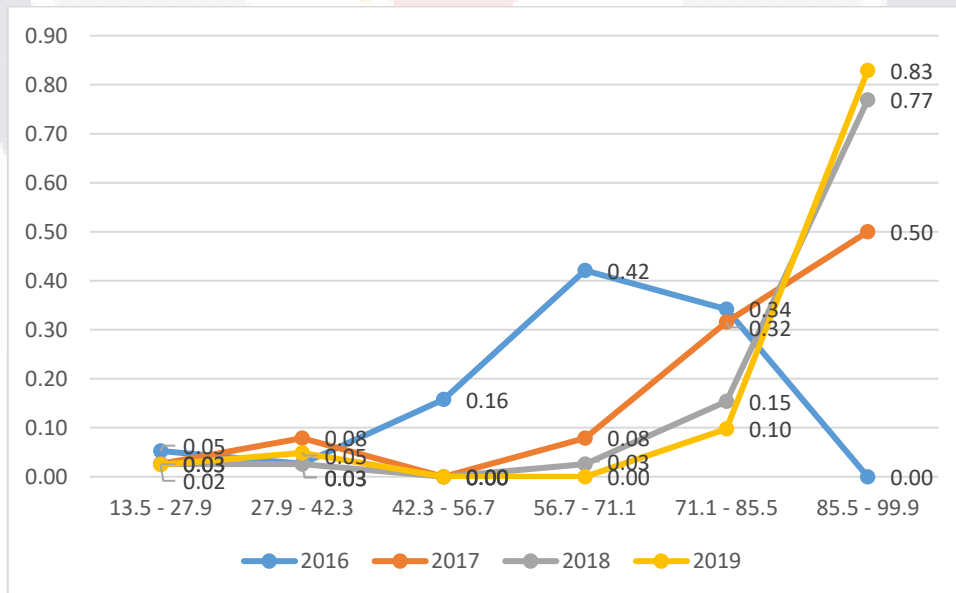


Figura 11. Distribución de oficinas con reentrevistas captadas por rango y año.

Fuente: Elaboración propia.

En la segunda iteración se realizó un análisis de asociación entre la cobertura de la reentrevista y las entrevistas logradas de la encuesta principal. Para esto, primero se verificó si los datos tienen una distribución normal, para identificar el tipo de prueba estadística que se debía aplicar. Primero se analizó el tipo de distribución.

Hipótesis: H0 = Los datos se comportan de forma normal.

Ha = Los datos no se comportan de forma normal.

La tabla 13 muestra la asimetría y curtosis de las dos variables en los periodos 2016 a 2019. En la variable de cobertura, los datos presentaron una asimetría negativa en todos los periodos y la curtosis presentó una gran concentración de valores en la región central de la distribución (leptocúrtica). La variable porcentaje de entrevistas logradas también presentó una asimetría negativa y una curtosis alta en 2017. Para los otros periodos presentó una concentración cercana a lo normal (mesocúrtica).

*Tabla 13. Medidas de distribución de cobertura de reentrevistas y porcentaje de entrevistas logradas.*

| Periodo | Cobertura   |             |             |             | Porcentaje entrevistas logradas |             |             |             |
|---------|-------------|-------------|-------------|-------------|---------------------------------|-------------|-------------|-------------|
|         | Asimetría   |             | Curtosis    |             | Asimetría                       |             | Curtosis    |             |
|         | Estadístico | Desv. Error | Estadístico | Desv. Error | Estadístico                     | Desv. Error | Estadístico | Desv. Error |
| 2016    | -1.629      | 0.383       | 3.281       | 0.750       | -0.807                          | 0.383       | 0.632       | 0.750       |
| 2017    | -1.975      | 0.383       | 2.957       | 0.750       | -1.356                          | 0.383       | 2.357       | 0.750       |
| 2018    | -2.874      | 0.378       | 9.278       | 0.741       | -1.296                          | 0.378       | 1.293       | 0.741       |
| 2019    | -3.425      | 0.378       | 14.217      | 0.741       | -0.917                          | 0.378       | 0.696       | 0.741       |

Fuente: Elaboración propia.

Como la muestra es menor a 50, se utilizó la prueba de Shapiro-Wilk para verificar si los datos siguen una distribución normal con un nivel de confianza del 95%, con el total de oficinas como grados de libertad. La tabla 14 muestra que sólo el porcentaje de entrevistas logradas en 2016 tiene una distribución normal (significancia mayor a 0.05).

Tabla 14. Prueba de normalidad para las variables de cobertura de reentrevista y porcentaje de entrevistas logradas.

| Periodo | Variable | Shapiro-Wilk |    |       |
|---------|----------|--------------|----|-------|
|         |          | Estadístico  | gl | Sig.  |
| 2016    | cob_p    | 0.858        | 38 | 0.000 |
|         | ent_p    | 0.949        | 38 | 0.083 |
| 2017    | cob_p    | 0.696        | 38 | 0.000 |
|         | ent_p    | 0.893        | 38 | 0.002 |
| 2018    | cob_p    | 0.646        | 39 | 0.000 |
|         | ent_p    | 0.867        | 39 | 0.000 |
| 2019    | cob_p    | 0.592        | 39 | 0.000 |
|         | ent_p    | 0.935        | 39 | 0.026 |

Fuente: Elaboración propia.

Dado que la asimetría es distinta de cero, la curtosis distinta de uno y la prueba de Shapiro – Wilk es menor a 0.05 se rechazó la hipótesis nula.

Al no seguir los datos una distribución normal se utilizaron pruebas no paramétricas. La prueba que se realizó para medir la asociación entre las dos variables fue la correlación de Spearman.

Hipótesis: H0 = No existe correlación entre el porcentaje de cobertura de reentrevistas y el porcentaje de entrevistas logradas.

Ha = Existe correlación entre el porcentaje de cobertura de reentrevistas y el porcentaje de entrevistas logradas.

La tabla 15 presenta el coeficiente de correlación de Spearman por año. Los años 2016 y 2018 presentaron una correlación baja, ya que el coeficiente obtuvo valores menores a 0.4, mientras que en 2017 y 2019 la correlación es moderada con valores entre 0.4 y 0.5. Como el nivel de significación para todos los periodos es menor a 0.05 se rechazó la hipótesis nula.

Tabla 15. Correlaciones entre el porcentaje de cobertura de reentrevista y porcentaje de entrevistas logradas.

| Periodo |                 |       |                             | cob_p | ent_p  |
|---------|-----------------|-------|-----------------------------|-------|--------|
| 2016    | Rho de Spearman | cob_p | Coefficiente de correlación | 1.000 | .321*  |
|         |                 |       | Sig. (bilateral)            |       | 0.049  |
|         |                 |       | N                           | 38    | 38     |
| 2017    | Rho de Spearman | cob_p | Coefficiente de correlación | 1.000 | .409*  |
|         |                 |       | Sig. (bilateral)            |       | 0.011  |
|         |                 |       | N                           | 38    | 38     |
| 2018    | Rho de Spearman | cob_p | Coefficiente de correlación | 1.000 | .363*  |
|         |                 |       | Sig. (bilateral)            |       | 0.023  |
|         |                 |       | N                           | 39    | 39     |
| 2019    | Rho de Spearman | cob_p | Coefficiente de correlación | 1.000 | .522** |
|         |                 |       | Sig. (bilateral)            |       | 0.001  |
|         |                 |       | N                           | 39    | 39     |

Fuente: Elaboración propia.

Con los resultados obtenidos se llegó a la conclusión de que existe una correlación directa, positiva y significativa, con una intensidad que oscila entre los niveles medio-bajo y medio entre el porcentaje de cobertura de reentrevistas y el porcentaje de entrevistas logradas.

#### 5.4.2 Desempeño

Dado que la cobertura sólo muestra el grado de avance, pero no el resultado obtenido, a continuación, se describen los resultados específicos obtenidos.

La figura 12 muestra la tasa de variación de las reentrevistas logradas en el ámbito nacional respecto al año anterior a partir de 2017. A partir de 2018, cuando se aplicaron los cambios derivados de la reingeniería, la tasa positiva se mantiene, aunque en menor grado, logrando un aumento del 9.5% en 2017, 1.0% en 2018 y 0.6% en 2019.

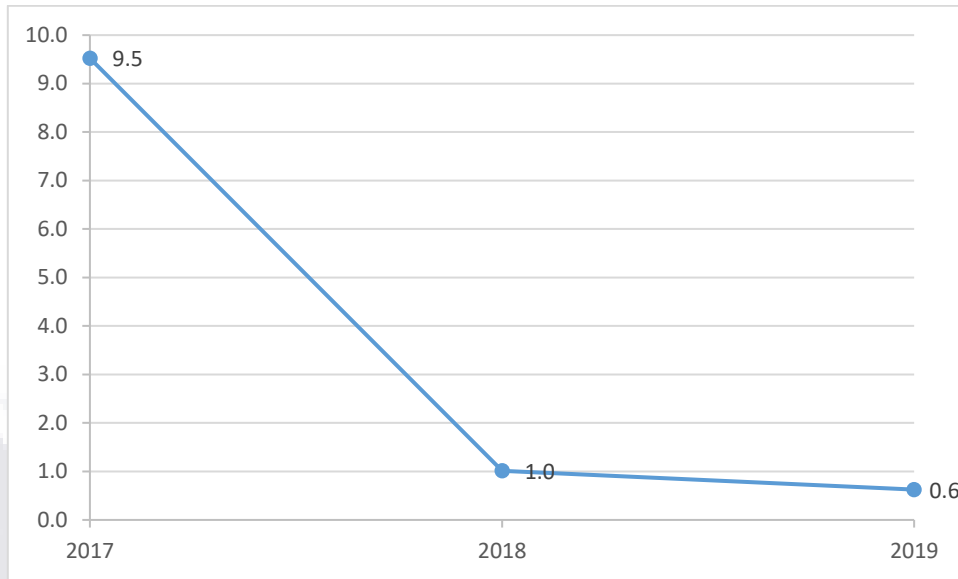


Figura 12. Tasa de variación de reentrevistas logradas por año en el ámbito nacional.

Fuente: Elaboración propia.

La tabla 16 muestra el análisis descriptivo de las reentrevistas logradas por año. Se puede ver que, de la misma manera que la tasa de variación de la figura 12, el mayor aumento en la media de reentrevistas logradas respecto al año anterior se logró en 2017, con un incremento del 7.6%. En los años posteriores se mantiene una tendencia positiva, con un aumento del 0.8% en 2018 y 1.0% en 2019.

Tabla 16. Estadísticos descriptivos de reentrevistas logradas en el ámbito nacional.

| Año  | Estadísticos descriptivos |                              |        |       |              |                                   |        |       |              |
|------|---------------------------|------------------------------|--------|-------|--------------|-----------------------------------|--------|-------|--------------|
|      | Total oficinas            | Total reentrevistas logradas |        |       |              | Porcentaje reentrevistas logradas |        |       |              |
|      |                           | Mínimo                       | Máximo | Media | Desv. típica | Mínimo                            | Máximo | Media | Desv. típica |
| 2016 | 38                        | 5                            | 748    | 316.6 | 166.3        | 24.3                              | 97.7   | 75.3  | 17.3         |
| 2017 | 38                        | 7                            | 1066   | 422.6 | 206.8        | 41.2                              | 95.8   | 82.9  | 11.4         |
| 2018 | 39                        | 14                           | 1160   | 458.9 | 215.9        | 47.5                              | 96.7   | 83.7  | 10.7         |
| 2019 | 40                        | 13                           | 1258   | 490.1 | 196.6        | 29.7                              | 96.1   | 84.7  | 12.3         |

Fuente: Elaboración propia.

Otro aspecto importante es el tipo de no respuesta. La figura 13 muestra una desagregación de la respuesta obtenida en las reentrevistas captadas por año en el ámbito nacional. Como se comentó en la figura 12, el cambio más significativo se presenta en 2017 respecto a 2016. A partir de 2018 continúa la tendencia positiva con una diferencia del 0.9% y de 0.5% en 2019. Respecto al tipo de no respuesta el código 03, que se refiere a que el informante se negó a dar la información, presenta una tendencia creciente anual, con un aumento respecto al año anterior del 6.6% en 2017, 6.7% en 2018 y 3.9% en 2019.

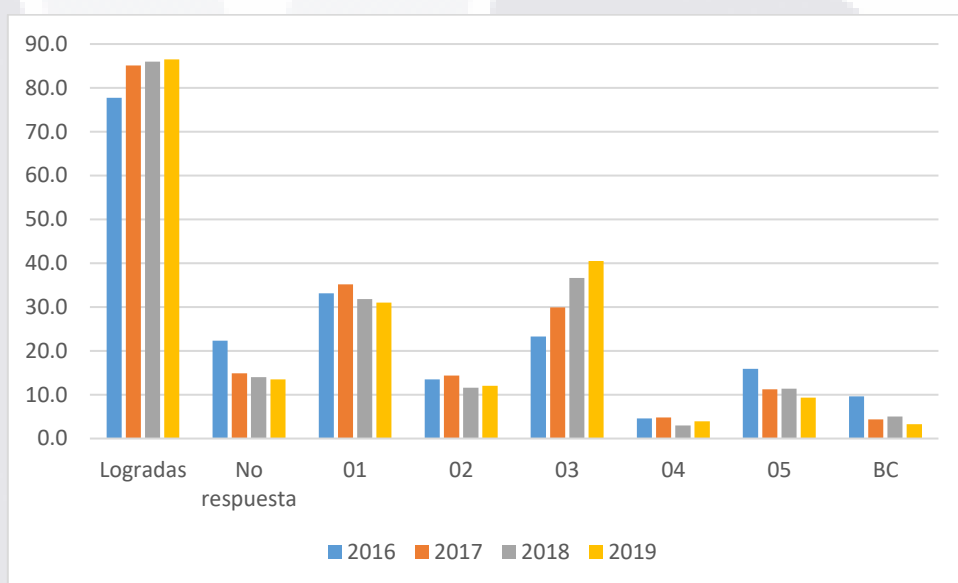


Figura 13. Porcentaje de reentrevistas por tipo de resultado y año.

Fuente: Elaboración propia.

La figura 14 presenta la tasa de variación de reentrevistas logradas por oficina y por año. Aunque la media nacional tuvo una tendencia positiva en 2019, con un incremento del 1.0%, como se puede observar en la tabla 16, el número de oficinas con tasa de variación negativa ha aumentado de 6 oficinas en 2017, a 13 oficinas en 2018 y 16 en 2019. Dos oficinas que se agregaron en 2019 no se incluyen en el análisis.

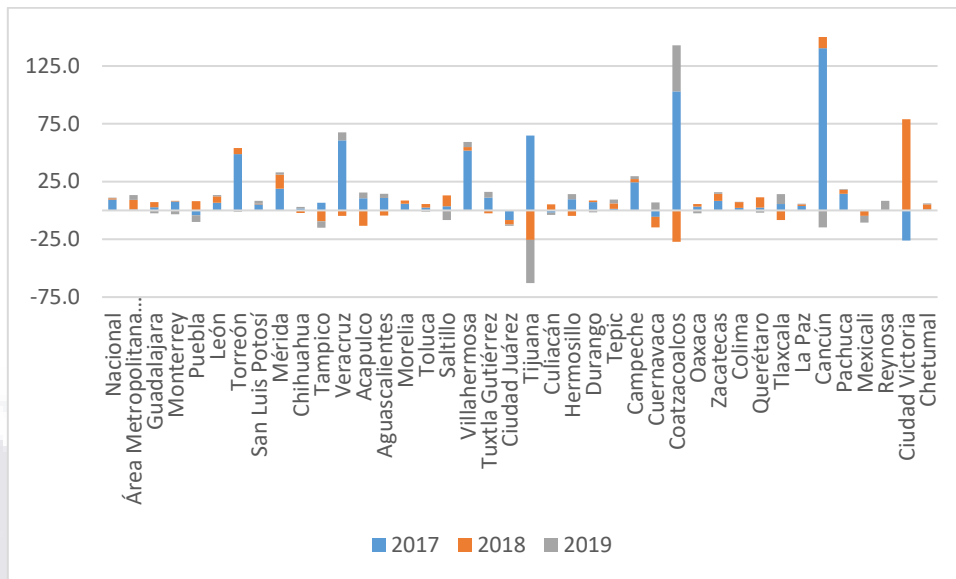


Figura 14. Tasa de variación de reentrevistas logradas por oficina y año.

Fuente: Elaboración propia.

La tabla 17 muestra la distribución de frecuencia del porcentaje de reentrevistas logradas por año. En las frecuencias se puede apreciar la tendencia de aumento del total de oficinas hacia la última clase.

Tabla 17. Tabla de distribución de frecuencia de reentrevistas logradas por año.

| LI           | LS   | 2016 |      | 2017 |      | 2018 |      | 2019 |      |
|--------------|------|------|------|------|------|------|------|------|------|
|              |      | F    | FR   | F    | FR   | F    | FR   | F    | FR   |
| 24.3         | 36.6 | 1    | 0.03 | 0    | 0.00 | 0    | 0.00 | 1    | 0.03 |
| 36.6         | 48.9 | 4    | 0.11 | 1    | 0.03 | 1    | 0.03 | 0    | 0.00 |
| 48.9         | 61.2 | 1    | 0.03 | 1    | 0.03 | 0    | 0.00 | 1    | 0.03 |
| 61.2         | 73.5 | 3    | 0.08 | 4    | 0.11 | 5    | 0.13 | 3    | 0.08 |
| 73.5         | 85.8 | 19   | 0.50 | 12   | 0.32 | 13   | 0.33 | 10   | 0.26 |
| 85.8         | 98.1 | 10   | 0.26 | 20   | 0.53 | 20   | 0.51 | 23   | 0.61 |
| <b>Total</b> |      | 38   | 1    | 38   | 1    | 39   | 1    | 38   | 1    |

Fuente: Elaboración propia.

La figura 15 muestra de manera gráfica el comportamiento del porcentaje de reentrevistas logradas por las oficinas operativas. Como se mencionó



anteriormente a partir de 2017 el 50% de las oficinas obtuvieron respuesta de al menos un 85.5%. En 2018 y 2019 continúa la tendencia positiva.

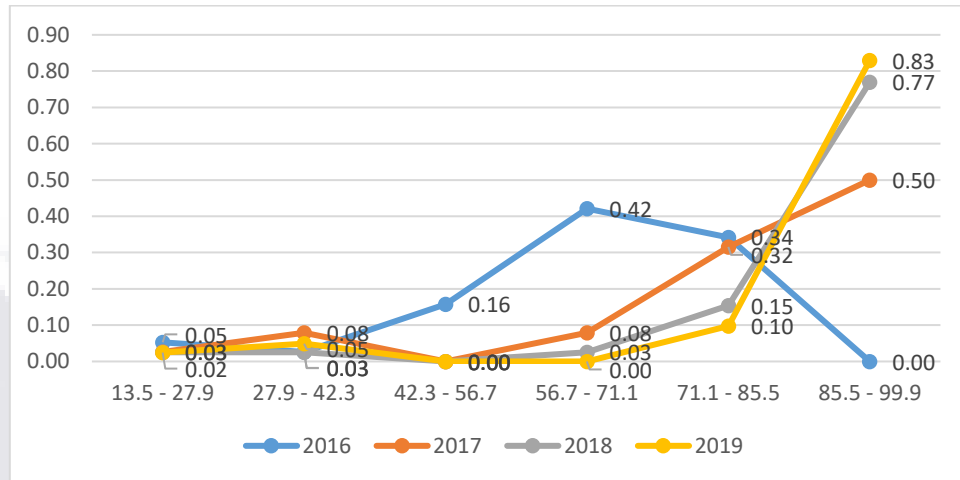


Figura 15. Distribución de oficinas con reentrevistas logradas por rango y año.

Fuente: Elaboración propia.

En la segunda iteración se realizó un análisis de asociación entre los resultados obtenidos de la reentrevista y los de la encuesta principal. Para esto también se verificó la distribución de la variable de porcentaje de reentrevistas logradas.

Hipótesis: H0 = Los datos se comportan de forma normal.

Ha = Los datos no se comportan de forma normal.

La tabla 18 muestra la asimetría y curtosis sólo de la variable de reentrevistas logradas, ya que la distribución de las entrevistas logradas se verificó anteriormente. Las reentrevistas logradas presentaron una asimetría negativa en todos los periodos y una concentración de valores cercana a la media en 2016 (mesocúrtica) y una concentración alta en los periodos posteriores (leptocúrtica).

Tabla 18. Medidas de distribución de reentrevistas logradas.

| Periodo | Reentrevistas logradas |             |             |             |
|---------|------------------------|-------------|-------------|-------------|
|         | Asimetría              |             | Curtosis    |             |
|         | Estadístico            | Desv. Error | Estadístico | Desv. Error |
| 2016    | -1.431                 | 0.383       | 1.425       | 0.750       |
| 2017    | -1.842                 | 0.383       | 4.017       | 0.750       |
| 2018    | -1.451                 | 0.378       | 2.265       | 0.741       |
| 2019    | -2.710                 | 0.383       | 9.415       | 0.750       |

Fuente: Elaboración propia.

La tabla 19 muestra la prueba de Shapiro – Wilk. Se puede apreciar que en ningún periodo el porcentaje de entrevistas logradas sigue una distribución normal (significancia mayor a 0.05).

Tabla 19. Prueba de normalidad para la variable porcentaje de reentrevistas logradas por año.

| Periodo | Variable | Shapiro-Wilk |    |       |
|---------|----------|--------------|----|-------|
|         |          | Estadístico  | gl | Sig.  |
| 2016    | res_00_p | 0.836        | 38 | 0.000 |
| 2017    | res_00_p | 0.825        | 38 | 0.000 |
| 2018    | res_00_p | 0.866        | 39 | 0.000 |
| 2019    | res_00_p | 0.726        | 38 | 0.000 |

Fuente: Elaboración propia.

Como la asimetría es distinta de cero, la curtosis distinta de uno y la prueba de Shapiro – Wilk tuvo un nivel de significancia menor a 0.05 se rechazó la hipótesis nula.

Una vez identificado que la variable no sigue una distribución normal, se verificó la asociación entre las dos variables mediante el coeficiente de correlación de Spearman.

Hipótesis: H0 = No existe correlación entre el porcentaje de reentrevistas logradas y el porcentaje de entrevistas logradas.

Ha = Existe correlación entre el porcentaje de reentrevistas logradas y el porcentaje de entrevistas logradas.

La tabla 20 presenta el coeficiente de correlación de Spearman por año. Los años 2016 y 2019 presentaron una intensidad de correlación media - baja, mientras que en 2018 la correlación es media. En 2017 no hay evidencia estadísticamente significativa que nos indique que existe una correlación, por lo que se rechaza la hipótesis nula para este periodo.

Tabla 20. Correlación entre reentrevistas y entrevistas logradas por año.

| Periodo |                 | res_00_p | ent_p                      |       |        |
|---------|-----------------|----------|----------------------------|-------|--------|
| 2016    | Rho de Spearman | res_00_p | Coeficiente de correlación | 1.000 | .392*  |
|         |                 |          | Sig. (bilateral)           |       | 0.015  |
|         |                 |          | N                          | 38    | 38     |
| 2017    | Rho de Spearman | res_00_p | Coeficiente de correlación | 1.000 | 0.238  |
|         |                 |          | Sig. (bilateral)           |       | 0.151  |
|         |                 |          | N                          | 38    | 38     |
| 2018    | Rho de Spearman | res_00_p | Coeficiente de correlación | 1.000 | .499** |
|         |                 |          | Sig. (bilateral)           |       | 0.001  |
|         |                 |          | N                          | 39    | 39     |
| 2019    | Rho de Spearman | res_00_p | Coeficiente de correlación | 1.000 | .374*  |
|         |                 |          | Sig. (bilateral)           |       | 0.021  |
|         |                 |          | N                          | 38    | 38     |

\*. La correlación es significativa en el nivel 0,05 (bilateral).

\*\*. La correlación es significativa en el nivel 0,01 (bilateral).

Fuente: Elaboración propia.

De acuerdo al coeficiente de correlación de Spearman la conclusión fue que existe una correlación baja, directa y ascendente entre las reentrevistas y las entrevistas logradas de los periodos 2016, 2018 y 2019, donde si aumenta el porcentaje de reentrevistas logradas puede aumentar el de las entrevistas logradas y viceversa.

Es importante mencionar que el coeficiente de correlación de Spearman sólo nos indica el grado de relación entre variables y no implica causalidad. Para tratar de

identificar el grado en el que la respuesta obtenida de la reentrevista puede afectar la respuesta de una entrevista se realizó un análisis de regresión lineal. Como variable independiente se utilizó el porcentaje de reentrevistas logradas y como variable dependiente el porcentaje de entrevistas logradas.

La tabla 21 muestra lo siguiente: el coeficiente de correlación de Pearson (R), el cual también indica una correlación media. El coeficiente de determinación R cuadrado es el porcentaje de variación de la variable dependiente que explica la variable independiente. En 2019 sólo el 23.8% de la variación en el porcentaje de entrevistas logradas se puede explicar por el porcentaje de reentrevistas logradas.

*Tabla 21. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas por año.*

| Periodo | R                 | R cuadrado | Error estándar de la estimación | Durbin-Watson |
|---------|-------------------|------------|---------------------------------|---------------|
| 2016    | .610 <sup>a</sup> | 0.372      | 3.2909                          | 1.474         |
| 2017    | .443 <sup>a</sup> | 0.196      | 3.7864                          | 1.152         |
| 2018    | .548 <sup>a</sup> | 0.300      | 3.3931                          | 1.382         |
| 2019    | .488 <sup>a</sup> | 0.238      | 2.9547                          | 1.528         |

a. Predictores: (Constante), res\_00\_p

b. Variable dependiente: ent\_p

Fuente: Elaboración propia.

La tabla 22 muestra los coeficientes de regresión parcial que definen la ecuación de regresión en puntuaciones directas. En todos los periodos los coeficientes son positivos y estadísticamente significativos (significancia menor a 0.05).

Tabla 22. Coeficientes del modelo de regresión lineal simple.

| Periodo | modelo      | Coeficientes no estandarizados |             | t      | Sig.  |
|---------|-------------|--------------------------------|-------------|--------|-------|
|         |             | B                              | Desv. Error |        |       |
| 2016    | (Constante) | 76.099                         | 2.421       | 31.435 | 0.000 |
|         | res_00_p    | 0.145                          | 0.031       | 4.616  | 0.000 |
| 2017    | (Constante) | 72.790                         | 4.568       | 15.936 | 0.000 |
|         | res_00_p    | 0.162                          | 0.055       | 2.963  | 0.005 |
| 2018    | (Constante) | 68.516                         | 4.327       | 15.836 | 0.000 |
|         | res_00_p    | 0.204                          | 0.051       | 3.983  | 0.000 |
| 2019    | (Constante) | 75.732                         | 3.299       | 22.955 | 0.000 |
|         | res_00_p    | 0.130                          | 0.039       | 3.354  | 0.002 |

Variable dependiente: ent\_p

Fuente: Elaboración propia.

El modelo de regresión lineal se construyó a partir de la siguiente ecuación:

$$y = a + bx$$

Donde  $a$  es una constante y  $b$  un coeficiente.

A manera de ejemplo, si en 2019 aumentara una unidad el porcentaje de reentrevistas logradas, el porcentaje de entrevistas logradas aumentaría 0.130.

Para validar si nuestro modelo es confiable, se deben verificar los siguientes supuestos: linealidad, independencia, homocedasticidad de varianzas y normalidad.

La figura 16 muestra el gráfico de dispersión de los residuos de 2019, en el cual se puede apreciar que los datos cumplen con el supuesto de linealidad, pero no el de homocedasticidad de varianzas, ya que se presentan problemas de dispersión irregular, donde la variabilidad de los errores aumenta conforme aumenta la media.

Sólo se presentó la gráfica de 2019, ya que en todos los periodos la distribución es similar.

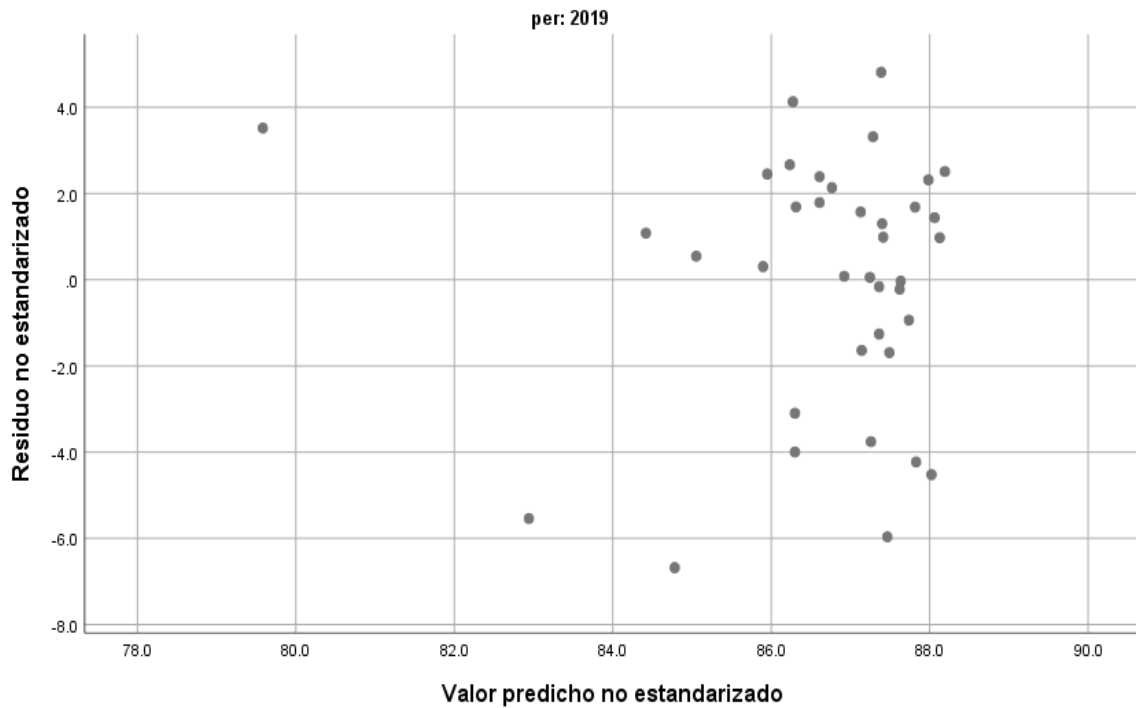


Figura 16. Gráfico de dispersión de residuos.

Fuente: Elaboración propia.

La independencia de los términos de error se refiere a que cada valor predicho es independiente y no afecta a otra predicción. La tabla 21 nos muestra el estadístico de Durbin-Watson, el cual mide el grado de autocorrelación entre el residuo (o error) correspondiente a cada observación y la anterior. Como los valores en todos los años son aproximadamente mayores a 1.4 la conclusión fue que los residuos son incorrelacionados, es decir cumplen el supuesto de independencia.

Por último, la tabla 23 muestra las pruebas de normalidad de los residuos, la significancia en ambas pruebas es menor a 0.05 para los periodos 2017 a 2019, por lo que los residuos no siguen una distribución normal.

Tabla 23. Pruebas de normalidad de residuos por año.

| per  | Kolmogorov-Smirnov <sup>a</sup> |    |       | Shapiro-Wilk |    |       |
|------|---------------------------------|----|-------|--------------|----|-------|
|      | Estadístico                     | gl | Sig.  | Estadístico  | gl | Sig.  |
| 2016 | 0.102                           | 38 | .200* | 0.971        | 38 | 0.407 |
| 2017 | 0.159                           | 38 | 0.016 | 0.889        | 38 | 0.001 |
| 2018 | 0.144                           | 39 | 0.040 | 0.919        | 39 | 0.008 |
| 2019 | 0.154                           | 38 | 0.023 | 0.935        | 38 | 0.030 |

Fuente: Elaboración propia.

Como los supuestos de normalidad y homogeneidad de varianzas no se cumplieron, el modelo de regresión no es adecuado para predecir la variación de los resultados de entrevista, pero sí se pudo utilizar para explicar de qué manera puede variar el resultado de entrevistas en función del porcentaje de reentrevistas logradas.

La tabla 24 muestra los resultados obtenidos con el modelo construido. Aunque el modelo no es adecuado, la diferencia entre el valor real y el valor predicho es mínima. De acuerdo al modelo, si en 2019 se hubiera logrado el 100% de reentrevistas la media de entrevistas logradas sería de 88.73%, en cambio, si el porcentaje de reentrevistas logradas fuera del 75% el valor obtenido para la variable dependiente es de 85.48%.

Tabla 24. Aplicación de modelo de regresión de entrevistas logradas en función de reentrevistas logradas.

| Periodo | a<br>(Constante) | b<br>(Coeficiente) | x (porcentaje de reentrevistas logradas) | y<br>(Predicción porcentaje de entrevistas logradas) | Porcentaje de entrevistas logradas | Diferencia predicción y valor real |
|---------|------------------|--------------------|--|--|------------------------------------|------------------------------------|
| 2016    | 76.099           | 0.145              | 75.3                                     | 87.02  | 87.00                              | -0.02                              |
| 2017    | 72.79            | 0.162              | 82.9                                     | 86.22  | 86.20                              | -0.02                              |
| 2018    | 68.516           | 0.204              | 83.70                                    | 85.59  | 85.61                              | 0.02                               |
| 2019    | 75.732           | 0.13               | 84.70                                    | 86.74  | 86.57                              | -0.17                              |

Fuente: Elaboración propia.



Cabe destacar que el modelo sólo explica la variación de un porcentaje que se encuentra entre el 20 y 30 por ciento. Como la correlación es directa, con una intensidad media-baja, pero significativa y ascendente se pudo concluir que la variación en el desempeño de las entrevistas logradas en función de las reentrevistas logradas existe. Para el caso de 2017, no hubo evidencia estadísticamente significativa para la afirmación anterior.

### 5.4.3 Visitas

El interés de analizar el número de visitas en un inicio era poder comparar el promedio de visitas necesarias para lograr una entrevista y una reentrevista. Al momento de integrar la información, se identificó que el número de visitas se registra para toda la muestra de la encuesta principal, pero no para la reentrevista.

El registro comienza a partir de 2017 y no presentó ajustes para la reingeniería, por lo que la variable se clasificó como una característica de la dinámica operativa y se realizaron pruebas para ver si es un factor que representa un impacto en los procesos.

La figura 17 muestra el comportamiento del promedio de visitas nacional por tipo de resultado. Se puede apreciar que la variación anual en cada tipo de resultado es casi nula.

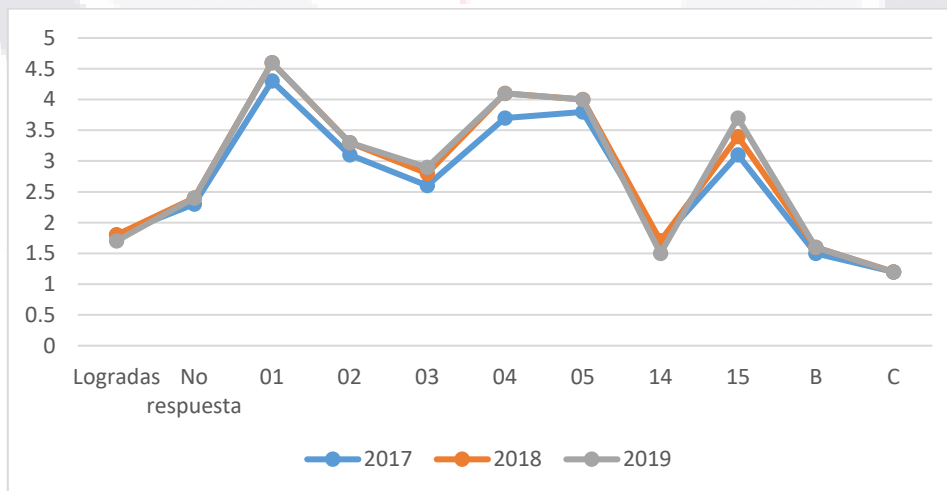


Figura 17. Promedio de visitas por tipo de respuesta y año.

Fuente: Elaboración propia.

Se realizó una prueba de normalidad para conocer la distribución de la variable promedio de visitas en el ámbito urbano.

Hipótesis: H0 = Los datos se comportan de forma normal.

Ha = Los datos no se comportan de forma normal.

La tabla 25 nos muestra los resultados obtenidos de la prueba de Shapiro-Wilk, la cual presentó un nivel de significación mayor a 0.05 por lo que se aceptó la hipótesis nula.

Tabla 25. Prueba de normalidad del promedio de visitas por año.

| Periodo | Shapiro-Wilk |    |       |
|---------|--------------|----|-------|
|         | Estadístico  | gl | Sig.  |
| 2017    | 0.949        | 33 | 0.126 |
| 2018    | 0.943        | 36 | 0.061 |
| 2019    | 0.949        | 36 | 0.099 |

Fuente: Elaboración propia.

Aunque el análisis descriptivo no presentó una diferencia significativa se agregó la variable promedio de visitas como variable independiente o predictor en el modelo de regresión lineal.

La tabla 26 muestra el coeficiente de correlación de Pearson (R) el cual se redujo aproximadamente a la mitad. El coeficiente de determinación R cuadrado ajustado nos muestra el porcentaje de variación que explican las variables independientes. Cabe mencionar que el porcentaje en los tres periodos no rebasa el 25%.

Tabla 26. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas y promedio de visitas por año.

| Periodo | R                 | R cuadrado | R cuadrado ajustado | Error estándar de la estimación | Durbin-Watson |
|---------|-------------------|------------|---------------------|---------------------------------|---------------|
| 2017    | .537 <sup>a</sup> | 0.288      | 0.241               | 3.4294                          | 0.891         |
| 2018    | .518 <sup>a</sup> | 0.268      | 0.224               | 3.4867                          | 1.307         |
| 2019    | .547 <sup>a</sup> | 0.300      | 0.257               | 2.8093                          | 1.325         |

a. Predictores: (Constante), vis\_00, res\_00\_p

b. Variable dependiente: ent\_p

Fuente: Elaboración propia.

La tabla 27 muestra los coeficientes de regresión parcial que definen la ecuación de regresión en puntuaciones directas. En todos los periodos los coeficientes de la variable de promedio de visitas logradas no son significativos con valores mayores a 0.05, por lo que no hubo evidencia estadística de que los coeficientes fueran adecuados para el modelo.

Tabla 27. Coeficientes del modelo de regresión lineal simple con dos variables independientes.

| Periodo | Coeficientes no estandarizados |        |             | t      | Sig.  |
|---------|--------------------------------|--------|-------------|--------|-------|
|         |                                | B      | Desv. Error |        |       |
| 2017    | (Constante)                    | 60.647 | 7.661       | 7.916  | 0.000 |
|         | res_00_p                       | 0.162  | 0.066       | 2.459  | 0.020 |
|         | vis_00                         | 6.761  | 3.916       | 1.727  | 0.095 |
| 2018    | (Constante)                    | 71.170 | 7.165       | 9.933  | 0.000 |
|         | res_00_p                       | 0.198  | 0.058       | 3.438  | 0.002 |
|         | vis_00                         | -1.187 | 3.630       | -0.327 | 0.746 |
| 2019    | (Constante)                    | 76.222 | 4.958       | 15.375 | 0.000 |
|         | res_00_p                       | 0.142  | 0.039       | 3.652  | 0.001 |
|         | vis_00                         | -0.828 | 2.745       | -0.302 | 0.765 |

a. Variable dependiente: ent\_p

Fuente: Elaboración propia.

Por último, para tratar de identificar algún cambio significativo en el comportamiento del promedio de visitas logradas se optó por realizar una comparación de medias utilizando la prueba t de Student.

La tabla 28 muestra los estadísticos descriptivos del promedio de visitas logradas por año. Tanto la media como la desviación típica son muy similares en los tres periodos.

Tabla 28. Estadísticos descriptivos de promedio de visitas logradas por año.

| Periodo | N  | Media | Desv. Desviación | Desv. Error promedio |
|---------|----|-------|------------------|----------------------|
| 2017    | 33 | 1.773 | 0.1606           | 0.0280               |
| 2018    | 36 | 1.794 | 0.1672           | 0.0279               |
| 2019    | 36 | 1.767 | 0.1821           | 0.0303               |

Fuente: Elaboración propia.

En base al análisis realizado de los estadísticos descriptivos de la tabla 28, el valor de la prueba de t de Student seleccionado fue de 1.8 visitas, por ser el promedio de visitas de los tres periodos. La prueba se llevó a cabo con un nivel de confianza del 95%.

Hipótesis: H0 = El promedio de visitas logradas es de 1.8.

Ha = El promedio de visitas logradas es distinto de 1.8

La tabla 29 muestra los resultados de la prueba t de Student para la variable promedio de entrevistas logradas con el valor de prueba seleccionado de 1.8 visitas para las pruebas de hipótesis. Como la significación bilateral es mayor a 0.05 se aceptó la hipótesis nula en los tres periodos, es decir hay evidencia estadísticamente significativa de que el promedio de visitas logradas en los tres periodos es de 1.8 visitas.

Tabla 29. Prueba t de Student para promedio de entrevistas logradas por año.

| Periodo | t      | gl | Sig. (bilateral) | Diferencia de medias | 95% de intervalo de confianza de la diferencia |          |
|---------|--------|----|------------------|----------------------|--|----------|
|         |        |    |                  |                      | Inferior                                       | Superior |
| 2017    | -0.975 | 32 | 0.337            | -0.0273              | -0.084   | 0.030    |
| 2018    | -0.199 | 35 | 0.843            | -0.0056              | -0.062   | 0.051    |
| 2019    | -1.099 | 35 | 0.279            | -0.0333              | -0.095   | 0.028    |

Fuente: Elaboración propia.

Al no presentar ajustes en los periodos 2017 a 2019 como parte del proceso de reingeniería y no presentar una diferencia significativa en el promedio de visitas, la conclusión fue que los cambios derivados del proceso de reingeniería no afectan el promedio de visitas de entrevistas logradas en la encuesta principal.

#### 5.4.4 Tiempo

De manera similar a la variable de visitas, el interés de analizar el tiempo promedio de entrevistas era para evaluar el impacto en el proceso de supervisión y después buscar una relación con la encuesta principal. Como el tiempo promedio se registra

en el control de visitas, de igual manera sólo se cuenta con información de la encuesta principal.

El registro comenzó a partir de 2017 y no presentó ajustes para la reingeniería, por lo que la variable se clasificó también como una característica de la encuesta principal y se hicieron pruebas para verificar que no sea un factor que represente un impacto derivado del proceso de supervisión.

La figura 18 muestra el tiempo promedio de las entrevistas logradas y no logradas en el ámbito nacional. Se puede apreciar un aumento en el tiempo promedio de entrevistas logradas en el año 2018 respecto al año anterior y en 2019 la diferencia no es significativa. En el tiempo promedio general y de entrevistas no logradas no se muestra una diferencia significativa.

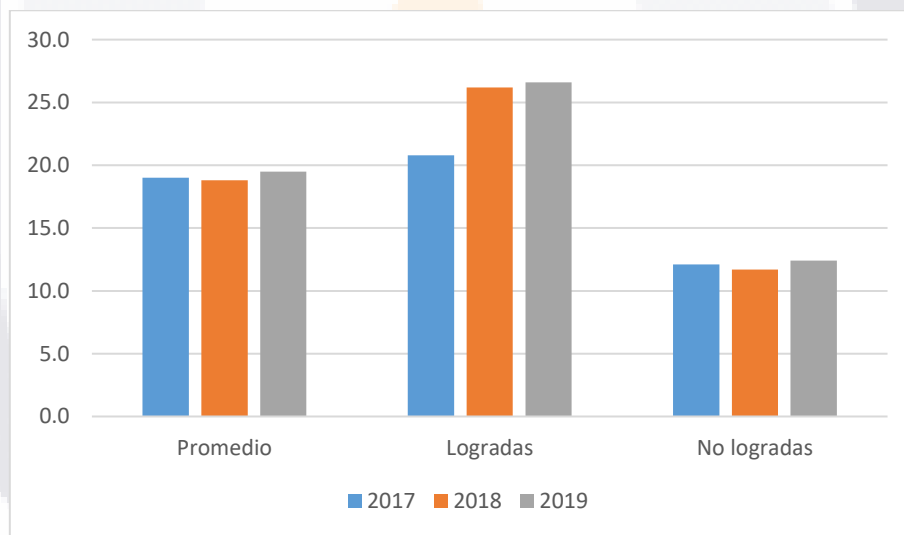


Figura 18. Tiempo promedio de entrevistas por año.

Fuente: Elaboración propia.

El aumento del tiempo promedio de una entrevista se puede dar por varios factores, como el aumento de preguntas en el instrumento de captación o un aumento en el número de residentes captados por vivienda o indicaciones operativas.

En este caso la diferencia en el promedio de entrevistas logradas se debe a que para 2017 sólo se tiene información del cuarto trimestre de 2017, por lo que sólo se

realizó el análisis con el promedio general para conocer si el registro de tiempo incide en los resultados de entrevista.

Primero se realizó una prueba de normalidad para conocer la distribución de la variable.

Hipótesis: H0 = Los datos se comportan de forma normal.

Ha = Los datos no se comportan de forma normal.

La tabla 30 muestra los resultados de la prueba de normalidad, como el estadístico de prueba Shapiro-Wilk presentó un nivel de significación mayor a 0.05 se aceptó la hipótesis nula.

*Tabla 30. Prueba de normalidad del tiempo promedio de entrevistas por año.*

| Periodo | Shapiro-Wilk |    |       |
|---------|--------------|----|-------|
|         | Estadístico  | gl | Sig.  |
| 2017    | 0.981        | 38 | 0.740 |
| 2018    | 0.979        | 39 | 0.679 |
| 2019    | 0.964        | 39 | 0.235 |

Fuente: Elaboración propia.

Al igual que el promedio de visitas el análisis descriptivo no presenta una diferencia significativa en el tiempo promedio de visitas. Con la finalidad de verificar si la variable es un factor que se refleje en los resultados se agregó el tiempo promedio como variable independiente o predictor en el modelo de regresión lineal.

La tabla 31 muestra el coeficiente de correlación de Pearson (R) el cual se redujo aproximadamente a la mitad. El coeficiente de determinación R cuadrado ajustado muestra el porcentaje de variación que explican las variables independientes, el cual tampoco rebasa el 26% en los tres periodos.

Tabla 31. Resumen del modelo de regresión lineal de entrevistas logradas en función de reentrevistas logradas y tiempo promedio de visitas por año

| Periodo | R                 | R cuadrado | R cuadrado ajustado | Error estándar de la estimación | Durbin-Watson |
|---------|-------------------|------------|---------------------|---------------------------------|---------------|
| 2017    | .471 <sup>a</sup> | 0.222      | 0.177               | 3.7783                          | 1.120         |
| 2018    | .551 <sup>a</sup> | 0.303      | 0.265               | 3.4319                          | 1.375         |
| 2019    | .488 <sup>a</sup> | 0.238      | 0.195               | 2.9966                          | 1.528         |

Fuente: Elaboración propia.

La tabla 32 muestra los coeficientes de regresión parcial que definen la ecuación de regresión en puntuaciones directas. En todos los periodos los coeficientes de la variable tiempo promedio de visitas no fueron significativos ya que obtuvieron valores mayores a 0.05, por lo que no hubo evidencia estadística de que los coeficientes fueran adecuados para el modelo.

Tabla 32. Coeficientes del modelo de regresión lineal simple con dos variables independientes.

| Periodo |                | Coeficientes no estandarizados |             | Coeficientes estandarizados | t      | Sig.  |
|---------|----------------|--------------------------------|-------------|-----------------------------|--------|-------|
|         |                | B                              | Desv. Error | Beta                        |        |       |
| 2017    | (Constante)    | 68.613                         | 5.990       |                             | 11.455 | 0.000 |
|         | res_00_p       | 0.174                          | 0.056       | 0.477                       | 3.128  | 0.004 |
|         | TiempoPromedio | 0.165                          | 0.153       | 0.164                       | 1.075  | 0.290 |
| 2018    | (Constante)    | 66.757                         | 6.138       |                             | 10.877 | 0.000 |
|         | res_00_p       | 0.210                          | 0.054       | 0.562                       | 3.917  | 0.000 |
|         | TiempoPromedio | 0.070                          | 0.171       | 0.059                       | 0.409  | 0.685 |
| 2019    | (Constante)    | 75.638                         | 4.899       |                             | 15.440 | 0.000 |
|         | res_00_p       | 0.130                          | 0.041       | 0.489                       | 3.198  | 0.003 |
|         | TiempoPromedio | 0.004                          | 0.137       | 0.004                       | 0.026  | 0.979 |

a. Variable dependiente: ent\_p

Fuente: Elaboración propia.

Para identificar algún cambio significativo en el comportamiento del tiempo promedio de visitas se realizó una comparación de medias utilizando la prueba t de Student.



La tabla 33 muestra los estadísticos descriptivos del tiempo promedio de visitas por año. Tanto la media como la desviación típica son muy similares en los tres periodos.

Tabla 33. Estadísticos descriptivos de tiempo promedio de visitas.

| Periodo | N  | Media  | Desv. Desviación | Desv. Error promedio |
|---------|----|--------|------------------|----------------------|
| 2017    | 38 | 19.074 | 4.1496           | 0.6732               |
| 2018    | 39 | 18.738 | 3.3646           | 0.5388               |
| 2019    | 39 | 19.395 | 3.7582           | 0.6018               |

Fuente: Elaboración propia.

En base al análisis realizado de los estadísticos descriptivos de la tabla 33, el valor de la prueba de t de Student seleccionado fue de 19.1 minutos, por ser el promedio de tiempo de los tres periodos. La prueba se llevó a cabo con un nivel de confianza del 95%.

Hipótesis: H0 = El tiempo promedio de visitas es de 19.1 minutos

Ha = El tiempo promedio de visitas es distinto de 19.1 minutos

La tabla 34 muestra los resultados de la prueba t de Student para la variable tiempo promedio de visitas con un valor de prueba de 19.1 minutos para las pruebas de hipótesis. Como la significación bilateral es mayor a 0.05 se aceptó la hipótesis nula en los tres periodos, es decir hay evidencia estadísticamente significativa de que el promedio de visitas logradas en los tres periodos es de 19.1 minutos.

Tabla 34. Prueba t de Student para tiempo promedio de visitas por año.

| Periodo | t      | gl | Sig. (bilateral) | Diferencia de medias | 95% de intervalo de confianza de la diferencia |          |
|---------|--------|----|------------------|----------------------|--|----------|
|         |        |    |                  |                      | Inferior                                       | Superior |
| 2017    | -0.039 | 37 | 0.969            | -0.0263              | -1.390   | 1.338    |
| 2018    | -0.671 | 38 | 0.506            | -0.3615              | -1.452   | 0.729    |
| 2019    | 0.490  | 38 | 0.627            | 0.2949               | -0.923   | 1.513    |

Fuente: Elaboración propia.

De igual manera que el registro de visitas, al no presentar ajustes en los periodos 2017 a 2019 como parte del proceso de reingeniería y no presentar una diferencia significativa en el tiempo promedio de visitas, la conclusión fue que los cambios derivados del proceso de reingeniería no afectaron el tiempo promedio de visitas en la encuesta principal.



## Capítulo 6. Evaluación

### 6.1 Valoración de los objetivos propuestos y alcanzados

En el presente trabajo práctico se plantearon cuatro objetivos específicos que se encuentran en el capítulo dos, objetivos de la intervención, los cuáles permitieron alcanzar el objetivo general. A continuación, se describen los resultados obtenidos de cada uno de los objetivos:

**Objetivo general:** Se llevó a cabo el análisis del proceso de supervisión, mediante el cual se identificaron las variables y periodos que permitieron medir el impacto de los cambios realizados al proceso, y la relación con los resultados obtenidos en el proyecto “Encuestas en hogares”. Se detectaron áreas de oportunidad y se generaron propuestas de análisis y herramientas para apoyar el proceso de toma de decisiones.

**Objetivo específico número uno:** Debido a que el área responsable realiza el análisis y seguimiento continuo del proceso de supervisión, y, por ende, de la encuesta principal, existe bastante información al momento que permite realizar distintos tipos de análisis. La selección inicial de variables se basó en la posibilidad de poder realizar comparaciones entre el proceso de supervisión y la encuesta principal. Las variables de cobertura y desempeño permitieron realizar conclusiones y mostrar el grado de impacto que han tenido estas variables y que podrían tener en un futuro. Las variables de promedio de visitas y tiempo promedio de visitas permitieron realizar hipótesis para identificar el grado de posibles cambios en los periodos de referencia.

**Objetivo específico número dos:** La infraestructura tecnológica del Instituto permitió llevar a cabo el análisis en todas las etapas del trabajo práctico, el hardware es el apropiado para realizar análisis de manera local y remota, las redes de comunicaciones permiten consultar, generar y presentar resultados de manera óptima, y se cuenta con licenciamiento de software que permite realizar el

procesamiento y explotación de información en distintas plataformas. Todo esto dentro del conjunto de normas, políticas y procesos que forman parte del sistema de seguridad de la información para preservar la confidencialidad, integridad y disponibilidad de la información.

**Objetivo específico número tres:** Se presentó al área responsable del proceso los resultados obtenidos del análisis de impacto obteniendo una respuesta positiva. De manera adicional se utilizó la herramienta de análisis Microsoft PowerBI para mostrar algunos resultados. Esta herramienta se eligió para la presentación debido a que permite integrar los datos desde diferentes recursos y visualizar resultados de una manera intuitiva, para personal familiarizado con el paquete Office 365. También ofrece la posibilidad de generar análisis en tiempo real y generación de tableros de control multiplataforma.

**Objetivo específico número cuatro:** Las sugerencias presentadas a la organización son las siguientes:

- Como la información se integra de manera continua es viable generar indicadores de supervisión enfocados a medir el impacto de los resultados obtenidos entre distintos procesos, por ejemplo, la captación y supervisión.
- Con la información histórica que se tiene al momento, se puede comenzar a realizar modelos que expliquen o ayuden a predecir resultados en base a estrategias teóricas, que ayuden a la toma de decisiones.
- Se pueden generar modelos específicos que integren la información para realizar la explotación en tableros de control en herramientas de análisis y visualización de datos como PowerBI.
- Existen variables cualitativas que se utilizan en controles como la observación que pueden ser analizadas para identificar sesgos en las valoraciones asignadas.
- La aplicación de técnicas estadísticas es viable ya que la información se encuentra almacenada en un sistema manejador de base de datos. Los

análisis y pruebas estadísticas realizados en el presente trabajo se pueden programar y agregar como insumos a los reportes y/o aplicaciones existentes que se encuentran en el sitio web del área responsable, permitiendo automatizar la generación de resultados y apoyar la toma de decisiones.

## **6.2 Alcances y limitaciones de la intervención**

El objetivo del presente trabajo consistió en medir el impacto de cambios realizados en el proceso de supervisión del proyecto “Encuestas en hogares” durante periodos específicos. Se contó con información de dos años anteriores y dos años posteriores que permitieron realizar los análisis y pruebas estadísticas para identificar comportamientos y tendencias y poder así, realizar conclusiones.

Los resultados se centraron principalmente en el logro de objetivos de supervisión como la cobertura y el desempeño; aunque existen otras variables, como los resultados específicos de no entrevista, las comparaciones que se realizan entre las respuestas obtenidas de la reentrevista y las obtenidas en las entrevistas de la encuesta principal, las respuestas obtenidas de la observación, la desagregación o agregación de información de periodos distintos, entre otras, que pueden ser analizadas en futuros trabajos con el objetivo de medir la calidad en los procesos de captación de información estadística.

Respecto a la encuesta principal, existe información de años anteriores con la que se pueden generar modelos más precisos para realizar comparaciones y/o predicciones sobre comportamientos hipotéticos y contrastar con los valores históricos obtenidos.

## **6.3 Aportes a la organización**

El análisis realizado en el presente trabajo pretende mostrar a la organización que la toma de decisiones operativas, las cuáles siempre tienen el objetivo de control y mejora continua en el proceso de captación pueden apoyarse de los análisis de estadística inferencial para soportar la toma de decisiones, utilizando la

infraestructura con la que cuenta actualmente el Instituto. Este tipo de análisis se puede aplicar a cualquier proyecto de encuestas en hogares que se levanta de manera regular o cuenta con más de un evento, con el objetivo de mantener o mejorar la calidad de los procesos de generación de información estadística.

#### **6.4 Recomendaciones para desarrollos futuros**

Durante el desarrollo del trabajo práctico se identificaron variables que pueden ser analizadas para medir la calidad de los procesos. Como el enfoque de la evaluación fue el de logro de objetivos, muchas de estas variables no se analizaron, pero es viable replicar el análisis de acuerdo a los objetivos planteados por los responsables del proceso. El objetivo de la generación del modelo de regresión lineal fue explicativo, sin embargo, con la información disponible y las pruebas de campo que se realizan cuando se van a presentar cambios, principalmente de tipo conceptual, se pueden generar modelos predictivos adecuados que apoyen la toma de decisiones.

## Capítulo 7. Conclusiones

Al concluir el análisis se tiene claridad sobre la influencia de los controles de supervisión hacia los resultados de la encuesta principal. Los resultados obtenidos para periodos anuales se pueden replicar con periodos específicos de acuerdo a las necesidades específicas de cada proyecto. A manera de ejemplo, es posible aumentar la precisión identificando unidades de observación que puedan tener más de una visita, en periodos mensuales o trimestrales, y verificar la tasa de cambio de respuesta obtenida una vez que la vivienda ha sido seleccionada, para llevar a cabo un control específico de supervisión.

Con los análisis estadísticos descriptivos e inferenciales utilizados en este trabajo práctico se podría explicar un fenómeno particular y realizar predicciones con información de calidad que permitan llevar a cabo el control y supervisión de un proceso determinado con información de calidad y oportuna para apoyar la toma de decisiones.

Otra aplicación del conocimiento adquirido es la identificación de áreas de oportunidad que permitan llevar a cabo la mejora continua de los procesos, ya que con la experiencia se suele dar por hecho algunos comportamientos o resultados obtenidos. Con las distintas pruebas estadísticas se puede identificar si realmente existe alguna tendencia en el comportamiento de los datos y la intensidad, para poder diseñar procesos en base a métodos científicos.

El hecho de supervisar o controlar un porcentaje específico o una muestra de una población o proceso, frecuentemente abre la posibilidad de pensar que las acciones realizadas hacia la muestra, como puede ser la supervisión, o el levantamiento de encuestas adicionales o módulos, no debe afectar de manera significativa el desempeño de una encuesta principal o proceso. Sin embargo, esta idea está basada sólo en el hecho de que la muestra es un porcentaje menor, y con las pruebas estadísticas adecuadas se puede identificar y predecir si realmente hay un impacto y en qué medida, lo cual puede ser mayor o menor que el supuesto inicial.



TESIS TESIS TESIS TESIS TESIS

Los análisis realizados permiten confirmar que efectivamente hubo una mejora en el proceso de supervisión. Desde el punto de vista de logro de objetivos, los cuáles se analizaron con las variables de cobertura y desempeño de la reentrevista, el impacto es positivo y significativo en los periodos de referencia, logrando mantener la tendencia positiva en el último año. Aunque cabe mencionar que la mejora más significativa se dio en 2017, un año antes de que se aplicara la reingeniería del proceso.

Respecto al promedio de visitas y tiempo promedio de visitas, el hecho de llevar el registro es una mejora, ya que permite dar seguimiento y abre la posibilidad de distintos análisis.

Para efectos de los objetivos del trabajo práctico, se puede concluir que no hubo un impacto significativo en las variables seleccionadas en los periodos de referencia, derivado de los cambios realizados al proceso de supervisión. Tomando como base esta conclusión, la propuesta es implementar los análisis estadísticos descriptivos e inferenciales de las variables críticas, o que generen más valor al área, e implementar o adoptar una herramienta que permita realizar análisis mediante tableros de control para apoyar el proceso de toma de decisiones.

El periodo de transición por el que se dio origen a la necesidad del trabajo práctico es una respuesta a la necesidad de información oportuna y de calidad por parte de la sociedad en general, por lo que la adopción de nuevas metodologías, estándares y mejora o diseño de nuevos procesos cada vez es más común, no sólo en el Instituto, sino en cualquier organización que esté orientada a la generación o explotación de información.

Durante la elaboración del trabajo práctico se aprovechó la infraestructura del Instituto para implementar la metodología que permitió procesar y analizar la información. Sin embargo, en la actualidad existe una variedad creciente de tecnologías, lenguajes de programación, entornos de desarrollo y librerías de uso

libre que permiten aplicar pruebas estadísticas robustas y la visualización de datos de una manera eficiente.

En este trabajo práctico se aplicaron pruebas estadísticas que son de aplicación general para el análisis de información, por lo que cualquier empresa u organización de cualquier sector puede aplicar el conocimiento adquirido para atender alguna necesidad específica.



**Glosario**

- BAM**            Monitoreo de Actividades del Negocio (Business Activity Monitoring).
- BPM**            Gestión de Procesos de Negocios (Business Process Management).
- BPMN**          Modelo y Notación de Procesos de Negocio (Business Process Model and Notation).
- BPMS**          Plataforma para la Gestión de Procesos de Negocio (Busines Process Management Suite).
- CAD**            Comité de Ayuda al Desarrollo (Development Assistance Committee).
- CAI**            Entrevista asistida por computadora (Computer Asisted Interview)
- CAPI**          Entrevista personal asistida por computadora (Computer Assisted Personal Interview).
- CATI**          Entrevista telefónica asistida por computadora (Computer Assisted Telephone Interview).
- CAWI**          CAWI: Entrevista por internet asistida por computadora (Computer Aid Web Interview).
- CRISP-DM**    Cross Industry Standard Process for Data Mining.
- EFTA**          Asociación Europea de libre Cambio (Eurepan Free Trade Association)
- EPA**            Encuesta de Población Activa
- GSBPM**        Modelo Genérico del Proceso Estadístico (Generic Statistical Business Process Model).
- INEGI**         Instituto Nacional de Estadística y Geografía.

|              |  |
|--------------|--|
| <b>MPEG</b>  | Modelo de Proceso Estadístico y Geográfico.  |
| <b>ONE</b>   | Oficina Nacional de Estadística.   |
| <b>ONU</b>   | Organización de las Naciones Unidas.   |
| <b>PAPI</b>  | Entrevista asistida con cuestionario en papel (Paper and Pencil Interview).                              |
| <b>PDA</b>   | Asistente Digital Personal (Personal Digital Assistant).   |
| <b>PMBOK</b> | Project Management body of knowledge.  |
| <b>PMI</b>   | Project Management Institute.  |
| <b>SDMX</b>  | Estándar para el intercambio de datos y metadatos estadísticos (Statistical Data and Metadata Exchange). |
| <b>SNIEG</b> | Sistema Nacional de Información Estadística y Geográfica.  |
| <b>SLA</b>   | Acuerdo de nivel de servicio (Service Level Agreement).  |
| <b>SOA</b>   | Arquitectura Orientada a Servicios (Service Oriented Architecture)                                       |
| <b>UNECE</b> | Comisión Económica de la Naciones Unidas para Europa (United Nations Economic Commission for Europe).    |

## Bibliografía

- Benes, E., & Walsh, K. (2018). *19th ICLS implementation: National LFS practices and implementation plans*. Recuperado de [https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms\\_646789.pdf](https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_646789.pdf)
- Ben-Menahem, S. M., Shrestha, Y. R., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, 61(4), 66-83. doi:10.1177/0008125619862257
- BPM Forum. (2013). *Business Process Management Pocket Guide*. Recuperado de <https://go.bizzdesign.com/BPM-Pocket-Guide>
- Brown, L., Citro, C., House, C., Marton, K., & Mackie, C. (2014). The Past, Present, and Future of Federal Surveys: Observations from the Committee on National Statistics. *Joint Statistical Meetings 2014*. Boston, Massachusetts.
- CAD. (2010). *Glosario de los principales términos sobre evaluación y gestión basada en resultados*. Recuperado de <https://www.oecd.org/dac/evaluation/2754804.pdf>
- Canada, S. (2011). In -depth review of household survey methods. *Second meeting of the 2011/2012 Bureau*. Geneva, Switzerland: Economic Commission for Europe conference of european statisticians. Recuperado de [https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2011/4-Household\\_survey\\_methods.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2011/4-Household_survey_methods.pdf)
- CONEVAL. (2014). *Manual para el diseño y la construcción de indicadores*. Recuperado de [https://www.coneval.org.mx/Informes/Coordinacion/Publicaciones%20oficiales/MANUAL\\_PARA\\_EL\\_DISENO\\_Y\\_CONTRUCCION\\_DE\\_INDICADORES.pdf](https://www.coneval.org.mx/Informes/Coordinacion/Publicaciones%20oficiales/MANUAL_PARA_EL_DISENO_Y_CONTRUCCION_DE_INDICADORES.pdf)

Díaz, F. (2008). Gestión de procesos de negocio BPM (Business Process Management), TICs y crecimiento empresarial ¿Qué es BPM y cómo se articula con el crecimiento empresarial? *Revista Universidad Empresa*, 7(15), 151-176. Recuperado de <https://revistas.urosario.edu.co/index.php/empresa/article/viewFile/1061/959>

Hitpass, B. (2017). *Business Process Management (BPM) Fundamentos y conceptos de implementación*. Santiago, Chile: BHH Ltda.

INEGI. (2010). *Norma técnica para la generación de estadística básica*. Recuperado de <http://www.snieg.mx/contenidos/espanol/normatividad/tecnica/Norma%20Técnica%20para%20la%20Generación%20de%20Estadística%20Básica.pdf>

INEGI. (2010). *Proceso estándar para encuestas por muestreo*. Recuperado de [http://www.snieg.mx/contenidos/espanol/normatividad/doctos\\_genbasica/proc\\_estandar\\_encuestas.pdf](http://www.snieg.mx/contenidos/espanol/normatividad/doctos_genbasica/proc_estandar_encuestas.pdf)

INEGI. (2011). *Diseño de la muestra en proyectos de encuesta*. Recuperado de [https://www.snieg.mx/documentacionportal/normatividad/vigente/doctos\\_genbasica/muestra\\_encuesta.pdf](https://www.snieg.mx/documentacionportal/normatividad/vigente/doctos_genbasica/muestra_encuesta.pdf)

INEGI. (2015). *Política de Calidad Institucional*. Recuperado de <http://sc.inegi.org.mx/repositorioNormateca/Pcal.pdf>

INEGI. (2015). *Principios y Buenas Prácticas para las Actividades Estadísticas y Geográficas del SNIEG*. Recuperado de [https://snieg.mx/DocumentacionPortal/Normatividad/vigente/Principios\\_Buenas\\_Practicas\\_Es\\_Geo\\_SNIEG.pdf](https://snieg.mx/DocumentacionPortal/Normatividad/vigente/Principios_Buenas_Practicas_Es_Geo_SNIEG.pdf)

INEGI. (2017). *Modelo del Proceso Estadístico y Geográfico MPEG*. Recuperado de [http://intranet.inegi.org.mx/calidad/wp-content/uploads/2017/03/Modelo\\_Procesos\\_INEGI\\_v1.0.pdf](http://intranet.inegi.org.mx/calidad/wp-content/uploads/2017/03/Modelo_Procesos_INEGI_v1.0.pdf)

INEGI. (2018). *Documento metodológico de supervisión*. Recuperado de <http://logistica.inegi.gob.mx/>

INEGI. (2018). *Reglamento interior del Instituto Nacional de Estadística y Geografía*. Recuperado de [http://sc.inegi.org.mx/repositorioNormateca/RIINEGI\\_18.pdf](http://sc.inegi.org.mx/repositorioNormateca/RIINEGI_18.pdf)

INEGI. (2019). *Filosofía institucional*. Recuperado de <http://intranet.inegi.org.mx/IdentidadInstitucional/Filosofia%20institucional/SitePages/default.aspx>

INEGI. (2019). *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*. Recuperado de [http://sc.inegi.org.mx/repositorioNormateca/O\\_13May19.pdf](http://sc.inegi.org.mx/repositorioNormateca/O_13May19.pdf)

INEGI. (2019). *Quiénes somos*. Recuperado de <https://www.inegi.org.mx/inegi/contenido/infoest.html>

International Organization for Standardization (ISO). (2015). *ISO 9001:2015(es) Sistemas de gestión de la calidad - Requisitos*. Recuperado de <https://www.iso.org/obp/ui/es/#iso:std:iso:9001:ed-5:v1:es>

ISO 9001. (2013). *ISO 9001 calidad. Sistemas de Gestión de Calidad según ISO 9001*. Recuperado de <http://iso9001calidad.com>

Meyer, B., Mok, W., & Sullivan, J. (2015). Household Surveys in Crisis. *Journal of Economic Perspectives*, 29(4), 199-226. doi:10.1257/jep.29.4.199

Organización de las Naciones Unidas (ONU). (2014). *Principios Fundamentales de las Estadísticas Oficiales*. Recuperado de [https://unstats.un.org/unsd/dnss/hb/S-fundamental%20principles\\_A4-WEB.pdf](https://unstats.un.org/unsd/dnss/hb/S-fundamental%20principles_A4-WEB.pdf)

Peñaloza, M. (2010). Teoría de las decisiones. *Perspectivas*, 13(25), 227-240.  
Recuperado de <http://www.redalyc.org/articulo.oa?id=425942454012>

Project Management Institute. (2013). *Guía de los fundamentos para la dirección de proyectos (Guía del PMBOK)* (5a. ed.).

Quevedo, V., Chía, J., & Rodríguez, A. (2014). *Midiendo el impacto*. Recuperado de <http://www.oei.es/salactsi/Cuba.pdf>

Real Academia Española. (2018). *Diccionario de la lengua española* (23.a ed.).  
Recuperado de <https://www.rae.es/>

Simon, H. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

Smart Vision Europe. (2018). *What is the CRISP-DM methodology?* Recuperado de <http://sv-europe.com/crisp-dm-methodology>

Statistical Working Group. (2018). *SDMX Glossary*. Recuperado de [https://sdmx.org/wp-content/uploads/SDMX\\_Glossary\\_Version\\_2\\_0\\_October\\_2018.docx](https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_2_0_October_2018.docx)

UNECE. (2013). *Generic Statistical Business Process Model GSBPM*. Recuperado de <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0>