



Gobierno Regional Junín



RESOLUCIÓN GERENCIAL GENERAL REGIONAL

N° 177 -2026-GRJ/GGR

Huancayo, 04 JUN. 2026

EL GERENTE GENERAL REGIONAL DEL GOBIERNO REGIONAL DE JUNÍN

VISTO:

Con Reporte N° 162-2026-GRJ/GRPPAT de fecha 21 de mayo del 2026; Reporte N° 184-2026-GRJ/GRPPAT/SGCTP de fecha 21 de mayo del 2026; Memorando N° 683-2026-GRJ-GGR/ORDITI de fecha 28 de abril del 2026; Memorando N° 656-2026-GRJ/GGR/ORDITI de fecha 22 de abril del 2026; Informe Legal N° 175-2026-GRJ-ORAJ, de fecha 20 de abril de 2026; Memorando N° 595-2026-GRJ-GRI, de fecha 13 de abril de 2026; Informe Técnico N° 006-2026-GRJ/GRPPAT/SGCTP de fecha 09 de abril del 2026; Memorando N° 456-2026-GRJ/GGR/ORDITI de fecha 26 de marzo del 2026; y demás documentos a los que hace referencia en la presente resolución; y,

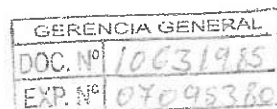
CONSIDERANDO:

Que, de conformidad con el artículo 191° de nuestra Carta Magna, concordado con el artículo 2° de la Ley N° 27867, Ley Orgánica de Gobiernos Regionales, los Gobiernos Regionales emanan de la voluntad popular, son personas jurídicas de derecho público, con autonomía política, económica y administrativa en los asuntos de su competencia, con la facultad de ejercer actos de gobierno, administrativos y de administración, sujetos al ordenamiento jurídico;

Que, mediante Memorando N° 595-2026-GRJ/GGR/ORDITI de fecha 13 de abril del 2026, por el Director Regional de Desarrollo Institucional y Tecnología de la Información, se remite el expediente que contiene la propuesta denominada "Guías Metodológicas para la Recolección y Validación de Datos en el Gobierno Regional de Junín", así como los Informes Técnicos N° 005-2026-GRJ/GRPPAT/SGCTP, emitidos por la Subgerencia de Cooperación Técnica y Planeamiento;

Que, de acuerdo con el Informe Técnico N° 006-2026-GRJ/GRPPAT/SGCTP de fecha 09 de abril del 2026 emitido por Sub Gerencia de Cooperación Técnica y Planeamiento, la propuesta ha sido formulada observando los criterios técnicos previstos en la Directiva N° 006-2011-GRJ-GGR/ORDITI y se encuentra alineada con el marco normativo vigente en materia de modernización de la gestión pública, gobierno digital, transparencia y protección de datos personales, emitiéndose opinión técnica favorable para su aprobación;

Que, mediante Informe Legal N° 175-2026-GRJ/ORAJ de fecha 20 de abril del 2026, la Oficina Regional de Asesoría Jurídica emite opinión legal sobre la viabilidad de aprobación del referido instrumento, precisando que no se advierte contravención al





Gobierno Regional Junín



ordenamiento jurídico vigente, así como que el mismo cuenta con sustento técnico emitido por el área competente;

Que, el artículo 6° de la Ley N° 27444 establece que todo acto administrativo debe estar debidamente motivado, expresando de manera clara los hechos relevantes y las razones jurídicas que sustentan la decisión;

Que, asimismo, se precisa que las **"Guías Metodológicas para la Recolección y Validación de Datos"** constituyen un instrumento de gestión interna de carácter orientador, cuya finalidad es establecer lineamientos técnicos para la adecuada recolección, tratamiento y uso de la información en el Gobierno Regional de Junín, no generando obligaciones jurídicas exigibles como las directivas, sino orientando la actuación de las unidades orgánicas en el marco de sus competencias;

Que, en ese sentido, resulta pertinente aprobar el referido instrumento técnico, a fin de fortalecer la gestión de la información institucional, promover la toma de decisiones basada en evidencia y contribuir al cumplimiento de los principios de eficiencia, transparencia y modernización de la gestión pública;

En uso de las facultades conferidas por la Ley N° 27867, Ley Orgánica de Gobiernos Regionales y sus modificatorias; y contando con la visación de la Oficina Regional de Desarrollo Institucional y Tecnología de la Información, la Subgerencia de Cooperación Técnica y Planeamiento, la Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial y la Oficina Regional de Asesoría Jurídica;

SE RESUELVE:

ARTÍCULO PRIMERO. – APROBAR las **"Guías Metodológicas para la Recolección y Validación de Datos en el Gobierno Regional de Junín"**, las mismas que como Anexo forman parte integrante de la presente Resolución.

ARTÍCULO SEGUNDO. – PRECISAR que las Guías aprobadas tienen naturaleza orientadora, constituyendo un instrumento técnico de apoyo para la gestión institucional, cuya aplicación se realizará en el marco de las competencias de cada unidad orgánica.

ARTÍCULO TERCERO. – DISPONER que la Subgerencia de Cooperación Técnica y Planeamiento, en coordinación con la Oficina Regional de Desarrollo Institucional y Tecnología de la Información, adopte las acciones necesarias para la implementación, difusión y seguimiento del instrumento aprobado.

ARTÍCULO CUARTO. – DISPONER la difusión de las Guías Metodológicas aprobadas a las dependencias y unidades orgánicas del Gobierno Regional de Junín, para su aplicación referencial en el desarrollo de sus funciones institucionales.





Gobierno Regional Junín



ARTÍCULO QUINTO. – DISPONER la publicación de la presente Resolución y sus anexos en los medios institucionales correspondientes.

ARTÍCULO SEXTO. – NOTIFICAR la presente Resolución a la Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial, a la Oficina Regional de Desarrollo Institucional y Tecnología de la Información, a la Subgerencia de Cooperación Técnica y Planeamiento y demás órganos competentes para los fines pertinentes.

REGÍSTRESE, COMUNÍQUESE Y CÚMPLASE.


CPC Mariamela Liz Quispe Salas
GERENTE GENERAL REGIONAL
GOBIERNO REGIONAL JUNÍN





GUÍA PARA VALIDACIÓN DE DATOS RECOLECTADOS

PLIEGO DEL GOBIERNO REGIONAL DE JUNÍN

2025



Subgerencia de Cooperación Técnica y
Planeamiento

GERENCIA REGIONAL DE PLANEAMIENTO,
PRESUPUESTO Y ACONDICIONAMIENTO TERRITORIAL



ELABORADO POR:	
VALIDADO POR:	
FECHA DE APROBACIÓN:	

**GUÍA PRÁCTICA PARA LA VALIDACIÓN DE DATOS EN EL GOBIERNO
REGIONAL DE JUNÍN**

ZOCIMO CÁRDENAS MUJE
Gobernador Regional

ECON. ANTONIA VEGA TERREL
Subgerente de Cooperación Técnica y Planeamiento

GOBIERNO REGIONAL DE JUNÍN

Gerencia Regional de Planeamiento, Presupuesto y
Acondicionamiento Territorial
Subgerencia de Cooperación Técnica y Planeamiento

GUÍA PRÁCTICA PARA LA **VALIDACIÓN DE DATOS** EN EL GOBIERNO REGIONAL DE JUNÍN

*Una guía amigable, práctica y fácil de entender
para todos los servidores del GORE Junín*

PRESENTACIÓN

El Gobierno Regional de Junín, a través de la Subgerencia de Cooperación Técnica y Planeamiento (SGCTP), presenta la "Guía Práctica para la Validación de Datos en el Gobierno Regional de Junín", un instrumento técnico y metodológico diseñado para fortalecer la gobernanza de datos institucionales.

En el contexto actual, la información constituye un activo estratégico fundamental para la gestión pública. La toma de decisiones informada, la planificación territorial y la rendición de cuentas dependen directamente de la calidad, veracidad y oportunidad de los datos que generan las distintas unidades orgánicas y ejecutoras de nuestra institución. Sin embargo, la experiencia nos indica que los datos frecuentemente presentan errores de digitalización, inconsistencias lógicas o falta de estandarización, lo que dificulta su uso y análisis.

Esta guía nace con el propósito de cerrar esas brechas. No busca ser un documento complejo de difícil comprensión, sino una herramienta amigable, práctica y fácil de entender para todos los servidores del GORE Junín. Aquí encontrará explicaciones sencillas sobre qué es la validación, por qué es crucial validar antes de procesar la información, y cuáles son los pasos exactos para asegurar que nuestros conjuntos de datos cumplan con los cinco criterios de calidad: exactitud, coherencia, completitud, actualidad y validez.

El documento establece un proceso de doble nivel de validación (en la Unidad Ejecutora y en la SGCTP), definiendo responsabilidades claras para cada actor involucrado, desde el personal de campo hasta los jefes de unidad. Asimismo, incorpora disposiciones para la protección de datos personales y la publicación bajo estándares de datos abiertos, alineando nuestra gestión con las normativas nacionales vigentes.

Invitamos a todo el personal a utilizar esta guía como referencia diaria en sus labores. Su aplicación permitirá que el Gobierno Regional de Junín cuente con información confiable, trazable y útil para el desarrollo de nuestra región.

Zocimo Cárdenas Muje
Gobernador Regional

SIGLAS

SIGLA	SIGNIFICADO
SGCTP	Subgerencia de Cooperación Técnica y Planeamiento
UE	Unidad Ejecutora
ORDITI	Oficina Regional de Desarrollo Institucional y Tecnologías de Información
INEI	Instituto Nacional de Estadística e Informática
RENIEC	Registro Nacional de Identificación y Estado Civil
Ministerio de Salud	Ministerio de Salud
Ministerio de Educación	Ministerio de Educación
SERVIR	Autoridad Nacional del Servicio Civil
DNI	Documento Nacional de Identidad
CSV	Valores separados por comas (Valores Separados por Comas - Formato de archivo)
JSON	Notación de objetos JavaScript (Notación de Objetos JavaScript - Formato de archivo)
XLSX	Hoja de cálculo XML de Excel (Formato de hoja de cálculo)
SHP	Shapefile (Formato de archivo geoespacial)
GRJ	Gobierno Regional de Junín (Código institucional en conjuntos de datos)
DREJ	Dirección Regional de Educación de Junín (Ejemplo de código en conjuntos de datos)

CONTENIDO DE LA GUÍA

PRESENTACIÓN	5
CONTENIDO DE LA GUÍA	7
CAPÍTULO 1:¿QUÉ ES LA VALIDACIÓN DE DATOS?.....	12
1.1 Explicación Sencilla.....	12
1.2 ¿Por qué es importante validar los datos?.....	12
1.3 ¿Quiénes participan en la validación?	13
CAPÍTULO 2: LOS 5 CRITERIOS DE CALIDAD DE LOS DATOS	14
CAPÍTULO 3: PASO A PASO PARA VALIDAR DATOS	15
3.1 Validación de Nivel 1 (En tu Unidad Ejecutora)	15
3.2 Validación de Nivel 2 (En la SGCTP)	17
CAPÍTULO 4:ERRORES MÁS COMUNES Y CÓMO CORREGIRLOS	21
4.1 Clasificación de los errores según su gravedad	21
4.2 Cómo corregir los errores.....	22
CAPÍTULO 5:DISPOSICIONES FINALES	23
6.1 Disposiciones Complementarias	23
6.2 Disposiciones Transitorias.....	23

ASPECTOS GENERALES

1. Objetivo

El objetivo central de la presente directiva es establecer los lineamientos técnicos, metodológicos y administrativos para asegurar la calidad, integridad, coherencia y confiabilidad de los datos recolectados, procesados o generados por las unidades organizacionales del Gobierno Regional de Junín (GORE Junín).

Estos lineamientos garantizan que la información utilizada para la planificación, el seguimiento, la evaluación y la toma de decisiones públicas cumpla estándares nacionales e internacionales, evitando errores, duplicidades y distorsiones que puedan afectar la gestión institucional.

2. Finalidad

La finalidad de este documento es proporcionar un instrumento práctico, claro y accesible que oriente la validación de datos en todas las áreas del GORE Junín.

Este lineamiento estandariza el proceso de validación, establece responsabilidades, define mecanismos de trazabilidad y asegura que la información estadística y geoespacial: sea veraz, esté actualizada, se encuentre completa, sea coherente con sus fuentes y sea trazable en todas sus etapas.

Asimismo, busca convertir los principios de gobernanza de datos en acciones concretas, auditables y replicables, asegurando que la información institucional pueda ser publicada y utilizada de manera transparente y segura.

3. Alcance

Esta directiva es de aplicación obligatoria para todas las direcciones regionales, subdirecciones, gerencias regionales, unidades ejecutoras y áreas técnicas que generan, procesan o reportan datos del GORE Junín que tienen la responsabilidad de generar y validar datos estadísticos. Su cumplimiento garantiza homogeneidad, comparabilidad y trazabilidad en toda la información institucional del GORE Junín.

4. Definiciones Conceptuales

- 4.1. Dataset:** Conjunto estructurado de datos con registros (filas) y variables (columnas), acompañado de metadatos obligatorios que permiten su correcta interpretación.
- 4.2. Metadatos:** Información que describe los datos: qué significan, cómo fueron obtenidos, las reglas de calidad aplicadas y las restricciones de uso.

- 4.3. **Repositorio institucional de datos:** Infraestructura que almacena y resguarda datasets con control de versiones, integridad y disponibilidad.
- 4.4. **Datos abiertos:** Datos publicados bajo licencias que permiten su uso, reutilización y distribución libre, siempre respetando la atribución correspondiente, según estándares de la Open Definition.
- 4.5. **Esquema de datos:** Estructura formal que especifica nombre, tipo, obligatoriedad y reglas de validación. Sirven como estructura base para validar cualquier dataset.
- 4.6. **Diccionario de datos:** Documento que detalla definiciones, reglas semánticas, ejemplos y condiciones para interpretar correctamente cada variable.
- 4.7. **Calidad de datos:** Todo dataset debe ser evaluado rigurosamente bajo cinco criterios clave: exactitud (El dato debe ser lo más cercano posible al valor real), coherencia (El dato no debe contradecirse con otros datos relacionados), completitud (No deben faltar datos importantes), actualidad (El dato debe reflejar la realidad del período que dice representar) y validez (El dato debe cumplir con las reglas definidas en el esquema y el diccionario de datos).
- 4.8. **Trazabilidad y control de versiones:** Cada conjunto de datos debe ir acompañado de metadatos mínimos que incluyan: autor, unidad responsable, fecha de recolección, fecha de validación, versión (V01, V02...), número de registros, errores encontrados y estado. El Repositorio Institucional guardará esta información con una firma digital, creando un historial completo y a prueba de manipulaciones.
- 4.9. **Estandarización de formatos y catálogos:** Para que los datos de diferentes oficinas puedan "conversar" entre sí, todos debemos usar los mismos formatos y códigos. Se exige el uso de formatos abiertos como CSV, JSON, XLSX, GeoJSON o SHP. Es obligatorio usar catálogos de códigos oficiales (como el UBIGEO del INEI para distritos o los tipos de documento de RENIEC) y está prohibido inventar códigos propios o abreviaturas no autorizadas.
- 4.10. **Seguridad de la información y protección de datos personales:** El tratamiento de datos personales (nombres, DNI, teléfonos, direcciones) debe cumplir estrictamente con la Ley N° 29733. Antes de cualquier publicación, se deben aplicar técnicas de anonimización para proteger la identidad y privacidad de los ciudadanos. Esto no es opcional, es una obligación legal y ética.

- 4.11. Responsabilidad institucional por la validación:** Cada Unidad Ejecutora o unidad orgánica es la dueña y principal responsable de la calidad de sus datos. El Planificador o Estadístico de la unidad actúa como el primer filtro (validador de nivel 1). El Jefe de la Unidad Ejecutora es quien da la aprobación final con su firma, asumiendo la responsabilidad institucional por la información enviada.
- 4.12. Publicación bajo estándares de datos abiertos:** Solo los datasets que superen el 100% de las validaciones y no tengan errores críticos podrán ser publicados. La publicación se hará en formatos abiertos, con sus metadatos y diccionario de datos, y siguiendo los lineamientos del Portal Nacional de Datos Abiertos (datos.gob.pe).
- 4.13. Revisión y corrección post-publicación.** Si se detecta un error en un dato ya publicado, el procedimiento es claro: se retira inmediatamente, se notifica a los usuarios, se publica una versión corregida (ej. V02) y se mantiene la versión anterior con una etiqueta de "obsoleta" (deprecated), explicando siempre el motivo de la corrección. La transparencia es clave incluso al corregir errores.

5. Participantes del proceso

Esta sección define con precisión las responsabilidades asignadas a cada actor en las distintas etapas del ciclo de vida del dato, desde su generación hasta su publicación final, asegurando que cada uno comprenda su rol y contribución a la calidad de la información institucional.

ACTOR INVOLUCRADO	RESPONSABILIDADES CLAVE
Subgerencia de Cooperación Técnica y Planeamiento (SGCTP)	Dirige el proceso, emite lineamientos, ejecuta la validación de datos recolectados de nivel 2, administra el repositorio y aprueba su publicación.
Jefe de la Unidad Ejecutora	Garantiza calidad institucional final por la calidad de los datos generados en su unidad y aprueba formalmente el envío de cada conjunto de datos validado a la SGCTP mediante un acto administrativo o correo institucional firmado digitalmente.
Jefes de las Unidades de Planeamiento / Estadístico de las UE	Realizan validación de nivel 1. Son responsables de la calidad interna de los datos, certifican la revisión mediante el "Acta de Validación Primaria" y gestionan la corrección de observaciones emitidas por la SGCTP.

Planificadores / Estadísticos de las unidades ejecutoras	Realizan la validación operativa en oficina y campo. Verifican la coherencia de la información desde su origen, analizan los flujos de generación de datos y aplican los instrumentos de validación estandarizados.
Analista Estadístico (de la SGCTP)	Ejecuta la validación de Nivel 2 (automatizada y manual). Revisa la consistencia y metadatos, genera reportes de error, asigna identificadores únicos, gestiona el Observatorio Regional y prepara los datasets para su publicación final.
Oficina Regional de Desarrollo Institucional y Tecnologías de Información (ORDITI)	Brindar el soporte tecnológico para la implementación y mantenimiento del Observatorio Regional y el Portal de Datos Abiertos. Habilitar automatismos para la validación y preparar los archivos para su publicación.
Oficina de Comunicaciones / Especialista en comunicaciones	Diseña e implementa las estrategias de difusión y promoción de los datos institucionales. Asegura que la información estadística validada sea accesible y comprensible para la ciudadanía y otros grupos de interés.
Unidades Generadoras de Datos	Generan los datos de su competencia en los formatos y plazos establecidos. Aplican las fichas técnicas y diccionarios de datos definidos por la SGCTP, verifican la coherencia de su información y anonimizan datos sensibles antes de su envío.
Personal de campo y/o encuestadores(*)	Cumplen estrictamente con los protocolos y formatos aprobados para la recolección de datos. Garantizan la confidencialidad de la información y reportan cualquier incidencia que pueda afectar la calidad de los datos recolectados.

(*) Solo en casos en que se apliquen instrumentos de campo (Encuestas, entrevistas).

CAPÍTULO 1:

¿QUÉ ES LA VALIDACIÓN DE DATOS?

1.1 Explicación Sencilla

La validación de datos es como revisar tu tarea antes de entregarla. Imagina que eres un profesor y debes verificar que las notas de tus alumnos estén bien escritas, que no falte ninguna, que los nombres estén correctos y que las sumas den el resultado esperado. Eso es exactamente lo que hacemos cuando validamos datos: nos aseguramos de que todo esté bien antes de usar esa información para tomar decisiones importantes.

En el Gobierno Regional de Junín, manejamos mucha información: datos de población, de salud, de educación, de obras, de proyectos, entre muchos otros. Si estos datos tienen errores, las decisiones que tomamos pueden ser equivocadas. Por eso es tan importante validar: para estar seguros de que lo que registramos es verdad y sirve para mejorar nuestra región.

1.2 ¿Por qué es importante validar los datos?

Validar los datos es importante por muchas razones. Aquí te explicamos las principales:

RAZÓN	EXPLICACIÓN SENCILLA
Para tomar buenas decisiones	Si los datos son correctos, las autoridades pueden decidir correctamente dónde construir un hospital, qué lugares necesitan más apoyo, o cuánto dinero destinar a cada proyecto.
Para generar confianza	Cuando los ciudadanos saben que la información del gobierno es confiable, confían más en las instituciones. La transparencia empieza con datos bien validados.
Para no repetir trabajo	Si encontramos errores a tiempo, no tenemos que volver a recolectar los datos. Ahorramos tiempo, dinero y esfuerzo.
Para cumplir con la ley	Existen normas que nos obligan a tener datos correctos y proteger la información personal de los ciudadanos. Validar nos ayuda a cumplir estas normas.
Para poder publicar	Los datos que publicamos en el Portal de Datos Abiertos deben ser de calidad. Solo así otros pueden usarlos y confiar en nuestra información.

1.3 ¿Quiénes participan en la validación?

La validación no es trabajo de una sola persona. Es un equipo donde cada uno tiene un rol importante. Piensa en una cadena: si un eslabón falla, toda la cadena se debilita. En la validación de datos pasa lo mismo: todos somos importantes.

¿QUIÉN?	¿QUÉ HACE?
Tú (el que recolecta)	Eres la primera persona que ve los datos. Tu trabajo es revisar que lo que anotes o registres esté bien desde el principio. Como cuando llenas un formulario: revisas que todo esté completo antes de enviarlo.
Planificador o Estadístico de tu unidad	Es como un supervisor que revisa el trabajo de todos. Agrupa los datos de tu área y hace una primera revisión completa. Si encuentra algo raro, te pregunta para corregirlo.
Jefe de tu Unidad Ejecutora	Es el responsable final. Él o ella firma diciendo que los datos de su unidad están bien. Es como el director que aprueba las notas finales.
Analista de la SGCTP	Es un especialista que usa herramientas especiales (programas de computadora) para revisar los datos de todas las unidades. Busca errores que a simple vista no se ven.
Subgerente de Cooperación Técnica	Dirige todo el proceso y aprueba la publicación final de los datos. Es como el árbitro que da el resultado oficial.

CAPÍTULO 2:

LOS 5 CRITERIOS DE CALIDAD DE LOS DATOS

Para saber si un dato es bueno, debemos revisar cinco aspectos importantes. Piensa en estos criterios como cinco preguntas que te haces para saber si el dato sirve o no. Cada pregunta es como un filtro: si el dato pasa los cinco filtros, entonces es un dato de calidad.

LOS 5 CRITERIOS DE CALIDAD

Un dato es de calidad cuando cumple todas estas condiciones:

- 1. EXACTITUD** → El dato debe ser igual a la realidad
Ejemplo: Si en el documento dice que hay 150 estudiantes, en tu base de datos también debe decir 150, no 1500 ni 15.
- 2. COHERENCIA** → El dato debe tener sentido con los demás datos
Ejemplo: Si una persona tiene 5 años, no puede tener título universitario. Si ves estos datos juntos, algo está mal.
- 3. COMPLETITUD** → No deben faltar datos importantes
Ejemplo: Si registras beneficiarios, cada uno debe tener nombre, DNI, distrito y edad. Si falta alguno, el registro está incompleto.
- 4. ACTUALIDAD** → El dato debe ser del período correcto
Ejemplo: Si el reporte es de marzo 2025, no puedes poner datos de diciembre 2024. Deben ser datos del período que corresponde.
- 5. VALIDEZ** → El dato debe seguir las reglas establecidas
Ejemplo: Si el formato de fecha debe ser DD/MM/AAAA, no puedes escribir "marzo 2025" ni "03-2025". Debes seguir el formato que se indicó.

CAPÍTULO 3:

PASO A PASO PARA VALIDAR DATOS

Ahora que ya sabes qué es la validación y por qué es importante, te explicamos cómo hacerlo. El proceso tiene dos niveles: primero se valida en tu unidad (Nivel 1) y luego en la Subgerencia de Cooperación Técnica y Planeamiento (Nivel 2). Aquí te explicamos cada nivel de forma sencilla.

3.1 Validación de Nivel 1 (En tu Unidad Ejecutora)

Este primer nivel es el más importante porque es donde se detecta la mayoría de los errores. Aquí el planificador o estadístico de tu unidad revisa los datos antes de enviarlos. Es como cuando revisas tu trabajo antes de entregarlo al profesor. Sigue estos pasos:

Paso 1: Ir al origen de los datos

Lo primero que debes hacer es ir a donde se crean los datos. No te quedes solo en la oficina: ve al lugar donde se registra la información. Esto te ayudará a entender cómo se generan los datos y detectar posibles problemas desde el inicio.

- **Visitar el origen:** Ir a la oficina o al lugar donde se registran los datos (un centro de salud, centro de producción, una obra, una escuela) para observar el proceso.
- **Identificar responsables:** Conversar con las personas que llenan los formatos y registran la información día a día.
- **Mapear el flujo:** Dibujar el camino que sigue el dato desde que nace hasta que llega a sus manos, para detectar cuellos de botella o puntos de error.
- **Análisis histórico:** Si es la primera vez que se valida un dato, investigar cómo se ha recolectado en el pasado para entender su contexto.
- **Revisión de encuestas:** Si los datos provienen de encuestas, se debe revisar el cuestionario, la metodología de muestreo, las evidencias de campo y el software usado para procesar las respuestas.

Paso 2: Organizar y documentar

Una vez que entiendes de dónde vienen los datos, hay que organizarlos y dejar registro de lo que hiciste. Esto es muy importante porque si alguien pregunta después, tendrás todo documentado.

- **Usa los formatos estándar que te da la SGCTP:** no inventes tu propio formato. Si todos usamos el mismo, es más fácil comparar.
- **Llena la Matriz de Validación:** es un Excel donde escribes la versión del archivo, la fecha, cuántos registros hay, qué errores encontraste y cómo los corregiste.

- **Guarda todas las evidencias:** fotos, actas, capturas de pantalla, GPS. Todo lo que demuestre que los datos son reales.

Paso 3: Revisar la estructura de los datos

Ahora toca revisar el archivo de datos propiamente dicho. Abre el Excel o el archivo CSV y verifica lo siguiente:

¿QUÉ REVISAR?	¿CÓMO LO HAGO?
Columnas completas	Mira que no falte ninguna columna que debería estar. Si el formato dice que debe haber 10 columnas, deben estar las 10.
Tipos de datos correctos	Los números deben ser números, las fechas deben ser fechas, el texto debe ser texto. A veces Excel confunde y pone texto donde debería haber números.
Valores permitidos	Si el campo es 'sexo', solo debe decir 'M' o 'F'. Si dice 'masculino' o 'hombre', está mal.
Formatos coherentes	Todas las fechas deben tener el mismo formato. Todos los números deben tener el mismo número de decimales.

Paso 4: Revisar la coherencia interna

Esta revisión es para ver si los datos 'tienen sentido' entre sí. A veces cada dato por separado está bien, pero cuando los juntamos, vemos que no pueden ser verdad. Revisa estas relaciones:

- **Edad y nivel educativo:** Si alguien tiene 5 años, no puede tener educación secundaria completa.
- **Fechas:** La fecha de nacimiento debe ser anterior a la fecha de inscripción.
- **Totales:** Si sumas los beneficiarios hombres y mujeres, debe dar el total de beneficiarios.
- **Ubicación:** El distrito debe pertenecer a la provincia que se indica, y la provincia debe pertenecer a Junín.

Paso 5: Preparar y enviar el paquete de datos

Ya revisaste todo. Ahora hay que preparar el paquete para enviarlo a la SGCTP. Este paquete debe estar completo, así no te lo devuelven por faltar algo. Incluye los siguientes elementos:

ELEMENTO	DESCRIPCIÓN
----------	-------------

Archivo de datos	El archivo Excel o CSV con los datos limpios y revisados.
Matriz de validación	El Excel donde registraste todos los errores encontrados y cómo los corrigiste.
Metadatos	Un archivo que explica qué significan los datos, quién los recolectó y cuándo.
Evidencias	Fotos, actas, capturas de GPS o cualquier documento que respalde los datos.
Acta de Validación Primaria	Un documento firmado en el que declaras que los datos fueron validados según las reglas.

Envía este paquete por correo electrónico institucional a la Subgerencia de Cooperación Técnica y Planeamiento. Asegúrate de que el jefe de tu unidad ejecutora firme el acta de validación primaria.

3.2 Validación de Nivel 2 (En la SGCTP)

Cuando tu unidad termina su revisión (Nivel 1), los datos llegan a la Subgerencia de Cooperación Técnica y Planeamiento para una segunda revisión más técnica. Allí, un analista estadístico usa herramientas especiales para encontrar errores que a simple vista no se ven. No te preocupes si te devuelven observaciones: es parte del proceso para asegurar la calidad.

PROCESO DE VALIDACIÓN DE NIVEL 2

PASO 1: Recepción y Registro

El analista de la SGCTP revisa que hayas enviado todo completo: archivo de datos, matriz, metadatos, evidencias y acta firmada. Si falta algo, te contacta para completarlo. Para ello, se revisa:

- ❖ Que la UE o unidad orgánica del GORE Junín haya enviado todo.
- ❖ Que los documentos estén completos
- ❖ Que la firma del responsable sea válida.

PASO 2: Validación Automatizada

Se usan programas de computadora que revisan miles de datos en segundos. Buscan duplicados, valores vacíos en campos obligatorios, valores fuera de rango y códigos que no existen en los catálogos oficiales. Para ello se revisa:

- ❖ Registros duplicados.

- ❖ Valores nulos en campos obligatorios.
- ❖ Valores fuera de rango.
- ❖ Códigos que no existen en los catálogos oficiales.
- ❖ Tipos de datos incorrectos (ej. texto en un campo numérico).

PASO 3: Validación Estadística y Lógica

El analista busca datos extraños (outliers), compara con datos históricos y con fuentes externas como el INEI. Si algo no cuadra, se consulta para verificar. Para ello se revisa:

- ❖ Atípicos estadísticos
- ❖ Distribuciones anómalas
- ❖ Comparación histórica
- ❖ Comparación con fuentes externas (INEI, MINSA, MINEDU, etc)
- ❖ Sólo para el caso de base de datos de Encuentra, no se realizan observaciones directas para evitar la manipulación, sino que se almacenan y se realiza una limpieza documentada para su posterior publicación

PASO 4: Validación Territorial

Se revisa que los distritos pertenezcan a las provincias correctas y que todo coincida con la geografía de Junín. Para ello se revisa:

- ❖ Coherencia con la realidad territorial.
- ❖ Consistencia entre provincias, distritos y unidades de análisis, dependiendo del espacio de recolección de datos.
- ❖ Identificación de anomalías que requieren verificación con la UE

PASO 5: Emisión del informe de validación

La emisión del informe será formulada por el analista estadístico de la SGCTP, el mismo que deberá incluir:

- ❖ Porcentaje de error
- ❖ Observaciones clasificadas (grave, media, leve)
- ❖ Recomendaciones técnicas.
- ❖ Nivel de cumplimiento
- ❖ Requerimiento de correcciones (si corresponde)

Este documento será ratificado mediante un reporte del Subgerente de la SGCTP y remitido de manera virtual o física. Este documento será oficial y quedará registrado.

PASO 6: Corrección y reenvío

Si el analista estadístico de la SGCTP o el subgerente encuentra errores, deberá seguir el siguiente procedimiento:

- ❖ Devuelve las observaciones a la UE o unidad orgánica.
- ❖ La UE o unidad orgánica corrige y reenvía.
- ❖ Se valida nuevamente la parte corregida.

Este documento es oficial y quedará registrado como versión.

PASO 7: Emisión de Informe

El analista prepara un informe con el porcentaje de error, las observaciones encontradas (clasificadas como graves, medias o leves) y las recomendaciones. Este informe te lo envían para que hagas las correcciones.

Una vez que el dataset está limpio y validado, cumpliendo con:

- ❖ Estructura correcta
- ❖ Lógica interna
- ❖ Coherencia documental
- ❖ Coherencia estadística
- ❖ Evidencias completas
- ❖ Metadatos válidos

El analista estadístico deberá asignar:

- ❖ Código institucional único
- ❖ Clasificación de acceso
- ❖ Versión aprobada

A partir de todos los procedimientos anteriores, el dataset puede ser usado institucionalmente y publicado.

Se consolida un informe final. El Jefe de la UE aprueba formalmente el dataset y la SGCTP le asigna un código único de identificación.

Entidad	Unidad orgánica / Unidad ejecutora	Año	Periodo de datos	Versión
GRJ	DREJ	2025	T1	V01
Gobierno Regional de Junín	Dirección Regional de Educación Junín		Primer Trimestre	Versión 01

- El código se desarrolla como: **GRJ-DREJ_2025Q1-V01.**

- La consolidación del informe final es pertinente, pero debería indicarse que dicho informe queda archivado en el repositorio con control de versiones.

FASE 6: Publicación o archivo

Con el dataset aprobado y codificado, se realizan los últimos pasos para su difusión:

1. **Anonimización:** El analista se asegura de que cualquier dato personal sensible sea eliminado o enmascarado (De corresponder o en casos específicos).
2. **Preparación:** La Oficina de Tecnologías de Información (ORDITI) prepara los archivos en formatos abiertos (CSV, JSON).
3. **Publicación:** El dataset se sube al repositorio estadístico y al Portal Nacional de Datos Abiertos (datos.gob.pe).
4. **Registro:** Se publica en el Repositorio Estadístico del Observatorio Regional de Junín.
5. **Difusión:** Se comunica la nueva publicación a través de la web, redes sociales y boletines.

¿Qué pasa si hay errores?

Si el analista encuentra errores, no te preocupes. Es normal que haya correcciones. El proceso es el siguiente:

1. Te envían las observaciones detalladas con el tipo de error y la fila donde está.
2. Tienes 3 días calendario para revisar y corregir los errores señalados.
3. Reenvías el archivo corregido con las evidencias de las correcciones.
4. El analista revisa nuevamente solo la parte corregida.

Aprobación Final y Publicación

Cuando los datos pasan todas las validaciones, se les asigna un código único de identificación. Por ejemplo: **GRJ-DREJ_2025Q1-V01** significa:

- GRJ = Gobierno Regional de Junín
- DREJ = Dirección Regional de Educación de Junín
- 2025Q1 = Año 2025, Primer trimestre
- V01 = Versión 01 del dataset

Después de la aprobación, los datos se preparan para publicarse en el Portal Nacional de Datos Abiertos (datos.gob.pe) y en el Observatorio Regional de Junín. Antes de publicar, se eliminan o se enmascaran los datos personales para proteger la privacidad de los ciudadanos.

CAPÍTULO 4:

ERRORES MÁS COMUNES Y CÓMO CORREGIRLOS

En este capítulo te mostramos los errores que aparecen más seguido cuando validamos datos. Conocerlos te ayudará a prevenirlos y a corregirlos rápidamente si aparecen.

TIPO DE ERROR	¿QUÉ SIGNIFICA?	EJEMPLO
Error de digitación	Cuando escribes un número o palabra diferente al original.	El documento dice 455 y escribes 4550.
Error de estructura	Cuando el formato del archivo está mal organizado.	Las columnas están desordenadas o faltan cabeceras.
Error de consistencia	Cuando dos datos que deberían coincidir no coinciden.	Nivel educativo 'inicial' con edad '30 años'.
Error de completud	Cuando faltan datos que son obligatorios.	El campo 'distrito' está vacío en varios registros.
Error de rango	Cuando un valor está fuera de lo permitido.	Edad = -1 o número de hijos = 45.
Error de duplicidad	Cuando el mismo registro aparece más de una vez.	La misma persona está registrada dos veces.
Error geográfico	Cuando la ubicación no corresponde.	El distrito no pertenece a la provincia indicada.
Error semántico	Cuando el significado está mal aunque el formato esté bien.	Código '01' usado como 'mujer' cuando el catálogo dice que es 'varón'.

4.1 Clasificación de los errores según su gravedad

No todos los errores son igual de graves. Algunos impiden que los datos se usen, mientras que otros se pueden corregir después. Aquí te explicamos los tres niveles:

NIVEL	IMPACTO	QUÉ HACER
GRAVE	El dato no sirve para nada. No se pueden tomar decisiones con él.	Corrección inmediata. El dataset no se aprueba hasta corregirlo.
MEDIO	Afecta varias variables y puede causar confusiones.	Corrección urgente, pero el dataset puede quedar en espera.
LEVE	No afecta el análisis principal, pero debería mejorarse.	Se registra y se puede avanzar. Se corrige cuando sea posible.

4.2 Cómo corregir los errores

Dependiendo del tipo de error, hay diferentes formas de corregirlo. Aquí te mostramos los métodos:

MÉTODO	CUANDO USARLO Y CÓMO
Corrección directa	Cuando el error es evidente y tienes la fuente que lo confirma. Simplemente corriges el dato en el archivo y registras el cambio en la matriz de validación.
Corrección por consulta	Cuando no estás seguro del valor correcto. Pides aclaración a la unidad ejecutora, adjuntas la evidencia y corriges solo cuando tengas la confirmación.
Corrección por nueva recolección	Cuando el dato no existe o fue perdido. Debes contactar a la fuente original, reconstruir los datos o repetir parte del proceso de recolección en campo.
Eliminación del registro	Solo para duplicados exactos, registros sin respaldo o datos imposibles de validar. Es el último recurso cuando no hay forma de recuperar la información.

CAPÍTULO 5:

DISPOSICIONES FINALES

6.1 Disposiciones Complementarias

1. **Primera disposición complementaria.** La Subgerencia de Cooperación Técnica y Planeamiento (SGCTP) publicará un informe anual sobre la calidad de los datos del GORE Junín en el Observatorio Regional.
2. **Segunda disposición complementaria.** La SGCTP es responsable de archivar y custodiar las copias de seguridad de todos los datasets validados en el repositorio institucional, bajo responsabilidad.
3. **Tercera disposición complementaria.** La Oficina Regional de Desarrollo Institucional y Tecnologías de Información (ORDITI), en coordinación con la SGCTP, implementará automatismos para la validación de integridad en los sistemas institucionales.
4. **Cuarta disposición complementaria.** Todo el personal encargado del acopio y validación de datos será capacitado semestralmente en técnicas y herramientas de validación.
5. **Quinta disposición complementaria.** Queda prohibido el uso de la información con fines comerciales por parte de cualquier servidor o funcionario involucrado, bajo sanción administrativa conforme a la Ley SERVIR.
6. **Sexta disposición complementaria.** La falsificación o el envío reiterado de datos no validados dará lugar a las acciones administrativas correspondientes según la Ley SERVIR.
7. **Séptima disposición complementaria.** Los casos no previstos en el presente lineamiento serán propuestos por los planificadores o estadísticos de las unidades ejecutoras y resueltos por la Subgerencia de Cooperación Técnica y Planeamiento del GORE Junín o quien haga sus veces.

6.2 Disposiciones Transitorias

8. **Primera disposición transitoria.** La presente directiva entra en vigencia a partir de su aprobación a cargo del Gerente General Regional y se mantendrá vigente hasta la derogatoria de la misma.
9. **Segunda disposición transitoria.** Las Unidades Ejecutoras / Unidades organizacionales que se encuentran acopiando información a la aprobación de este lineamiento adaptarán sus formatos a las directivas vigentes en un plazo de 12 meses.
10. **Tercera disposición transitoria.** Dentro de los 12 primeros meses, la Subgerencia de Cooperación Técnica y Planeamiento y la Oficina Regional de Desarrollo Institucional y Tecnologías de Información habilitarán scripts de validación automatizada para las principales fuentes de datos.

ANEXOS

- Anexo 1: Acta de validación primaria de datos
- Anexo 2: Hoja de validación institucional
- Anexo 3: checklist rápido de validación
- Anexo 4: Formato de trazabilidad integral
- Anexo 5: Formato para solicitud de correcciones
- Anexo 6: Conformidad de datos corregidos
- Anexo 7: Formato de control de publicación
- Anexo 8: Formato para control de catálogos de datos y códigos
- Anexo 9: Formato para documentar excepciones de datos
- Anexo 10: Formato de alertas temprana de calidad de datos
- Anexo 11: Guía de anonimización de datos personales
- Anexo 12: Formato de informe de calidad anual

Anexo 1
ACTA DE VALIDACIÓN PRIMARIA DE DATOS
(Para la Oficinas de Planeamiento / o la que haga de sus veces de las
Unidades Ejecutoras)

Este es el documento formal que la Unidad Ejecutora o unidad orgánica del GORE Junín debe llenar y firmar para certificar que ha realizado la Validación de Nivel 1. Es la "declaración jurada" de la calidad interna de los datos.

Estructura del Acta de Validación Primaria:

1. Identificación:

- Unidad Ejecutora
- Responsable de la UE
- Fecha
- Nombre del dataset
- Periodo evaluado

2. Objeto: Declarar que el dataset ha sido validado internamente.

3. Alcance del Proceso: Un checklist para marcar qué tipo de validaciones se aplicaron (en campo, automatizada, semántica, etc.).

ÍTEM	SÍ	NO	OBSERVACIONES
Dataset completo	☐	☐	
Metadatos completos	☐	☐	
Diccionario actualizado	☐	☐	
Evidencias	☐	☐	
Sin duplicados	☐	☐	
Sin nulos obligatorios	☐	☐	
Acta firmada	☐	☐	

4. Resultados: Una tabla resumen con los indicadores de calidad

Nº	Variable	Revisión estructura	Revisión lógica	Revisión documental	Observaciones	Cumple (S/N)

5. Declaratoria: Un texto donde la UE declara bajo responsabilidad la veracidad de los datos y el cumplimiento de la ley de protección de datos.

Declaro que la información contenida en este dataset ha sido revisada bajo los criterios de:

- ☒ Completitud
- ☒ Coherencia
- ☒ Consistencia
- ☒ Integridad
- ☒ Evidencias documentales
- ☒ Uso del instrumento vigente

Observaciones (si hubiera):

6. **Entrega de Paquete:** Lista de los archivos incluidos en el .zip.
7. **Conclusión:**
 - ☐ Dataset validado
 - ☐ Dataset validado con observaciones
 - ☐ Dataset no validado
8. **Firmas:** Espacio para las firmas del Responsable de Planeamiento o estadísta y del Jefe de la UE.

Anexo 2
HOJA DE VALIDACIÓN INSTITUCIONAL
(Subgerencia de Cooperación Técnica y Planeamiento – GORE Junín)

Este formato es de uso interno de la SGCTP. Es el acta que se genera después de la Validación de Nivel 2, donde se registra el resultado final: Aprobado, Aprobado con Observaciones o Rechazado.

INNFORME DE VALIDACIÓN PRIMARIA (EN CASO EXISTAN ERRORES)

1. Identificación del dataset
2. Unidad ejecutora
3. Métodos de revisión utilizados
4. Resultados generales

Nº	Tipo de validación	Resultado	Detalle / Evidencia	Observaciones	Nivel (Grave/Medio/Leve)
1	Validación automatizada				
2	Validación estadística				
3	Validación documental				
4	Validación lógica				
5	Validación territorial				

5. Detalle de inconsistencias
6. Acciones aplicadas
7. Conclusiones y recomendaciones
8. Nivel de cumplimiento:
 - ✓ Alto
 - ▲ Medio
 - ✗ Bajo

Firma del responsable.

INFORME TÉCNICO DE VALIDACIÓN AVANZADA – SGCTP

1. Datos generales del dataset
2. Procedimiento de validación

ÍTEM	SÍ	NO	OBSERVACIONES
Validación automatizada ejecutada	☐	☐	
Control de estructura	☐	☐	
Validación estadística	☐	☐	
Validación territorial	☐	☐	
Registro de errores	☐	☐	
Informe completado	☐	☐	

3. Resultados de validación automatizada
4. Resultados estadísticos
5. Resultados de validación territorial
6. Resultados documentales
7. Porcentajes de error
8. Observaciones clasificadas
9. Recomendaciones técnicas
10. Conclusión:
 - ☐ Aprobado
 - ☐ Aprobado con observaciones
 - ☐ No aprobado

Firma del analista.

ANEXO 3: CHECK LIST RÁPIDO DE VALIDACIÓN

Una guía práctica para el Analista de la SGCTP, para asegurar que no se olvide de revisar ningún punto crítico durante la Validación de Nivel 2. El contenido mínimo se establece en 6 partes, el mismo que deberá estar de la siguiente manera:

1. RECEPCIÓN Y AUTENTICIDAD DEL EXPEDIENTE

Verifica que el expediente recibido sea legítimo, completo y autorizado.

ÍTEM	SÍ	NO	OBSERVACIONES
1.1 El nombre del archivo coincide con el formato institucional	☐	☐	
1.2 El Acta de Validación (Nivel 1) está firmada por la UE	☐	☐	
1.3 Se adjunta el Dataset en el formato solicitado (CSV, XLSX, etc.)	☐	☐	
1.4 Se incluye el Diccionario de Datos	☐	☐	
1.5 Los metadatos están anexados	☐	☐	
1.6 Las evidencias son legibles y suficientes	☐	☐	

2. VALIDACIÓN SINTÁCTICA (AUTOMÁTICA)

Revisión estructural del dataset mediante herramientas automáticas (Excel, Python, R, scripts institucionales).

ÍTEM	SÍ	NO	OBSERVACIONES
2.1 No existen registros duplicados	☐	☐	
2.2 No hay campos obligatorios con valores nulos	☐	☐	
2.3 Todas las variables tienen el tipo de dato correcto	☐	☐	
2.4 No existen valores fuera de rango	☐	☐	
2.5 No hay errores de formato en fechas	☐	☐	
2.6 Todos los códigos pertenecen al catálogo institucional	☐	☐	

3. VALIDACIÓN SEMÁNTICA (MANUAL)

Revisión del significado, coherencia y contexto del dato.

ÍTEM	SÍ	NO	OBSERVACIONES
3.1 Los valores atípicos (outliers) están identificados	☐	☐	
3.2 Los valores atípicos están justificados o explicados	☐	☐	
3.3 Las fechas respetan la lógica del evento (no hay inconsistencias temporales)	☐	☐	
3.4 Las variables dependientes son coherentes (reglas de negocio)	☐	☐	
3.5 No hay conflictos entre variables condicionales	☐	☐	
3.6 Los valores reportados tienen sustento en evidencias	☐	☐	

4. METADATOS Y TRAZABILIDAD

Se asegura de que el dataset pueda ser reproducido, auditado y comprendido por cualquier área.

ÍTEM	SÍ	NO	OBSERVACIONES
4.1 Los metadatos están completos (origen, fecha, responsable, metodología)	☐	☐	
4.2 El Diccionario de Datos coincide con el dataset	☐	☐	
4.3 Las evidencias están ordenadas y etiquetadas	☐	☐	
4.4 Se adjunta la Matriz de Validación de Nivel 1	☐	☐	
4.5 Se adjunta la versión del dataset	☐	☐	
4.6 La trazabilidad (cambios y correcciones) está documentada	☐	☐	

5. PROTECCIÓN DE DATOS PERSONALES

Revisión obligatoria según normativa de privacidad y la Guía de Anonimización.

ÍTEM	SÍ	NO	OBSERVACIONES
5.1 No existen nombres y apellidos de personas	☐	☐	
5.2 No se incluyen DNI, correos personales o teléfonos	☐	☐	
5.3 Las direcciones han sido agregadas (no aparecen direcciones exactas)	☐	☐	
5.4 No existen datos sensibles sin anonimizar (salud, discapacidad, etc.)	☐	☐	
5.5 Las coordenadas han sido generalizadas a nivel distrito o sector	☐	☐	
5.6 El dataset final cumple la Guía de Anonimización	☐	☐	

6. DECISIÓN FINAL

El analista debe clasificar el resultado final de acuerdo a los hallazgos.

RESULTADO	MARQUE	MOTIVO / CONDICIÓN
APROBADO	☐	No existen errores graves. Los errores medios están corregidos.
APROBADO CON OBSERVACIONES	☐	Existen errores medios que no afectan el análisis, pero deben corregirse en la siguiente versión.
NO APROBADO	☐	Existen errores graves, datos no confiables o falta de documentación.

Firma del Analista SGCTP: _____

Fecha: ____ / ____ / ____

ANEXO 4:
FORMATO DE TRAZABILIDAD INTEGRAL

Etapas	Fecha	Responsable	Documento	Resultado	Versión

ANEXO 5:
FORMATO PARA SOLICITUD DE CORRECCIONES
Dirigido de la SGCTP hacia el centro de producción de datos (UE o unidad
orgánica del GORE Junín)
(Puede ser remitida por un correo de acuerdo al siguiente detalle)

SOLICITUD DE CORRECCIÓN DE DATA

Unidad Ejecutora: _____

Dataset: _____

Fecha: ____ / ____ / ____

Observaciones encontradas:

1. _____
2. _____
3. _____

Fecha límite para corregir: ____ / ____ / ____

Firma SGCTP.

ANEXO 6:
CONFORMIDAD DE DATOS CORREGIDOS

(La respuesta será remitida vía correo electrónico, por medio de un mensaje)

CONFORMIDAD DE DATOS CORREGIDOS

Declaro que las observaciones emitidas por SGCTP el ____ / ____ / ____ han sido atendidas.

- ☒ Todas
- ☒ Parcialmente
- ☒ No atendidas

Evidencia adjunta:

Firma UE.

ANEXO 7:
FORMATO DE CONTROL DE PUBLICACIÓN
(El control deberá ser llevado por el analista estadístico de la SGCTP en
coordinación con la Oficina de Comunicaciones)

Dataset	Estado (publicado/no publicado)	Fecha de aprobación	Fecha de publicación	Plataforma	Responsable

ANEXO 8:
FORMATO PARA CONTROL DE CATÁLOGOS DE DATOS Y CÓDIGOS
(El analista estadístico hará uso del control de catálogo de datos y códigos)

Variable	Catálogo usado	Versión	Aplicación correcta (S/N)	Inconsistencias detectadas	Acción

ANEXO 9:
FORMATO PARA DOCUMENTAR EXCEPCIONES DE DATOS
(Podrá ser usado por el estadista o planificador de la UE o la unidad orgánica
del GORE Junín / en casos excepcionales podrá usarlo el analista estadístico
de la SGCTP del GORE Junín)

Nº	Registro	Valor inusual	Regla que incumple	Justificación técnica	Responsable	Evidencia

ANEXO 10:
FORMATO DE ALERTAS TEMPRANA DE CALIDAD DE DATOS
(Este formato permite detectar riesgos antes de su publicación del dataset)

Nº	Riesgo detectado	Indicador	Umbral	Nivel	Responsable	Acción requerida

Ejemplos de alertas:

- Más del 10% de campos obligatorios faltantes
- Inconsistencias territoriales
- Variaciones atípicas respecto al periodo anterior

ANEXO 11:

GUÍA DE ANONIMIZACIÓN DE DATOS PERSONALES

1. TIPOS DE DATOS PERSONALES A PROTEGER

1.1. Datos personales directos

Identifican de inmediato a una persona:

- Nombres
- DNI
- Dirección exacta
- Teléfono
- Firma
- Fotografía

1.2 Datos personales sensibles

Tienen mayor nivel de protección:

- Salud
- Discapacidad
- Opinión política
- Origen étnico
- Religión
- Identidad de género
- Situación penal

1.3 Datos cuasi-identificadores

No identifican directamente, pero en combinación sí:

- Fecha de nacimiento
- Distrito de residencia
- Centro educativo
- Centro de salud
- Número de dependientes

2. PRINCIPIOS PARA LA ANONIMIZACIÓN

- **Minimización:** solo conservar variables indispensables.
- **Irreversibilidad:** impedir reconstruir la identidad.
- **Proporcionalidad:** anonimizar según riesgo.
- **Utilidad:** mantener la funcionalidad del análisis sin comprometer la privacidad.
- **Trazabilidad:** registrar cada transformación.

3. TÉCNICAS DE ANONIMIZACIÓN PERMITIDAS

Para el GORE Junín se autorizan las siguientes técnicas:

3.1. Eliminación de variables identificatorias

Eliminar columnas como:

- Nombre
- DNI
- Apellidos
- Dirección exacta

3.2. Agregación

Convertir valores específicos en categorías amplias.

Ejemplo:

- Edad exacta → rango de edad
- Distrito → provincia

3.3. Enmascaramiento (masking)

Reemplazar parte de la información.

Ejemplo: DNI → ***12345

3.4. Seudonimización

Reemplazar el identificador por un código interno.

OJO: No es anónimo si existe una tabla que permite revertir.

3.5. Perturbación estadística

Alterar ligeramente los datos sin perder su utilidad estadística.

3.6. Redondeo y truncamiento

Especialmente en datos numéricos de alto riesgo.

4. VARIABLES QUE DEBEN SER ANONIMIZADAS OBLIGATORIAMENTE

VARIABLE	ACCIÓN REQUERIDA
Nombres y apellidos	Eliminar oseudonimizar
DNI	Eliminar o enmascarar
Dirección	Agregar (ej. provincia)
Teléfono	Eliminar
Firma digital	Eliminar
Coordenadas exactas	Generalizar a nivel distrito o sector
Datos de salud	Anonimizar o agregar categorías

5. PROCEDIMIENTO PARA ANONIMIZAR (PASO A PASO)

a) Paso 1. Identificar datos personales y sensibles

Revisar diccionario, estructuración y evidencias.

b) Paso 2. Evaluar riesgo de reidentificación

Considerar: tamaño de muestra, unicidad, combinación de variables.

c) Paso 3. Seleccionar técnica adecuada

Eliminación → si el campo no aporta
Agregación → si se necesita mantener análisis
Seudonimización → si se requiere trazabilidad

d) Paso 4. Ejecutar la anonimización

Aplicar transformaciones en el dataset.

e) Paso 5. Registrar la transformación

Usar el formato institucional de trazabilidad.

f) Paso 6. Validar irreversibilidad

Asegurar que ya no pueda identificarse a nadie.

g) Paso 7. Aprobar y publicar

La SGCTP autoriza la publicación interna/externa.

6. MATRIZ DE ANONIMIZACIÓN

Variable	Dato personal	Técnica aplicada	Nivel de riesgo	Resultado	Evidencia	Responsable

7. ERRORES COMUNES A EVITAR

- Seudonimizar sin eliminar tabla de reversión
- Publicar coordenadas exactas
- Mezclar datos sensibles con identificadores
- No documentar el proceso

8. PRODUCTO FINAL

Un dataset que:

- ☒ No identifica a ninguna persona
- ☒ Mantiene su utilidad estadística
- ☒ Cumple normativa peruana
- ☒ Puede ser publicado sin riesgo

ANEXO 12: FORMATO DE INFORME DE CALIDAD ANUAL

Gobierno Regional de Junín – SGCTP

“Informe Institucional de Calidad del Dato – Año _____”

1. IDENTIFICACIÓN DEL INFORME

- Año evaluado: _____
- Área responsable: SGCTP
- Responsable técnico: _____
- Fecha de elaboración: ____ / ____ / ____

2. OBJETIVO DEL INFORME

Presentar los resultados del monitoreo anual de la calidad de los datos generados por las Unidades Ejecutoras del GORE Junín, identificando:

- avances,
- retrocesos,
- nivel de cumplimiento,
- errores recurrentes,
- y oportunidades de mejora.

3. METODOLOGÍA APLICADA

- Revisión documental
- Validación automatizada
- Validación estadística
- Control de calidad
- Trazabilidad de versiones
- Observación en campo (si aplica)

4. RESUMEN EJECUTIVO

Breve descripción de:

- estado general de la calidad del dato
- principales hallazgos
- tendencias observadas
- riesgos identificados

5. CUMPLIMIENTO POR UNIDAD EJECUTORA

UE / UO GORE Junín	Dataset evaluados	% de errores graves	% de errores medios	% de errores leves	Nivel de cumplimiento
--------------------------	----------------------	---------------------------	---------------------------	--------------------------	--------------------------

--	--	--	--	--	--

6. PRINCIPALES ERRORES DETECTADOS

Tipo de error	Frecuencia	Tendencia vs año previo	Impacto	Observaciones

7. INDICADORES ANUALES DE CALIDAD

Indicador	Fórmula	Resultado	Meta institucional	Variación
% de campos completos				
% duplicados				
% inconsistencias				
% datasets aprobados				
Tiempo promedio de validación				

8. HALLAZGOS DESTACADOS

Incluye:

- Buenas prácticas
- Problemas estructurales
- Brechas de capacidades
- Riesgos para la planificación

9. RECOMENDACIONES

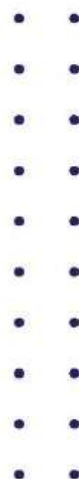
- Correctivas
- Preventivas
- De capacitación
- De actualización tecnológica
- De estandarización

10. PLAN DE MEJORA

Acción	Responsable	Plazo	Indicador asociado	Estado

11. ANEXOS

- Listado de datasets revisados
- Reportes de validación
- Matriz de errores consolidada
- Evidencias documentales
- Gráficos comparativos



GUÍA PARA LA RECOLECCIÓN DE DATOS

PLIEGO DEL GOBIERNO REGIONAL DE JUNÍN

2025



• Subgerencia de Cooperación Técnica y
• Planeamiento

GERENCIA REGIONAL DE PLANEAMIENTO,
PRESUPUESTO Y ACONDICIONAMIENTO TERRITORIAL



ELABORADO POR:	
VALIDADO POR:	
FECHA DE APROBACIÓN:	

GUÍA PRÁCTICA PARA LA RECOLECCIÓN DE DATOS EN EL GOBIERNO REGIONAL DE JUNÍN

ZOCIMO CÁRDENAS MUJE
Gobernador Regional

ECON. ANTONIA VEGA TERREL
Subgerente de Cooperación Técnica y Planeamiento

Copyright © 2025
Derechos Reservados

- Informe técnico emitido por la Oficina Regional de Desarrollo Institucional y Tecnologías de Información.
- Revisado por el equipo técnico de la Subgerencia de Cooperación Técnica y Planeamiento.

Dirección: Jr. Loreto N° 363, Huancayo
Oficina de Enlace en Lima: Jirón Gregorio Paredes N° 316
Lima Cercado - Teléfono: (01)3723753

GOBIERNO REGIONAL DE JUNÍN

Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial

Subgerencia de Cooperación Técnica y Planeamiento

GUÍA PRÁCTICA PARA LA RECOLECCIÓN DE DATOS EN EL GOBIERNO REGIONAL DE JUNÍN

Una guía amigable, práctica y fácil de entender

para todos los servidores del GORE Junín

CONTENIDO DE LA GUÍA

PRESENTACIÓN	7
CAPÍTULO 1:.....	8
ASPECTOS GENERALES	8
CAPÍTULO 1: ¿QUÉ ES LA RECOLECCIÓN DE DATOS?	19
1.1. Definición:.....	19
1.2. ¿Por qué es importante recolectar datos bien?.....	19
1.3. Las 6 Fases del Proceso de Recolección	20
CAPÍTULO 2: TIPOS DE DATOS Y CÓMO RECOLECTARLOS.....	21
2.1. Datos Estructurados vs No Estructurados	21
2.2. Fuentes de Datos: Primarias y Secundarias	21
2.3. Métodos para Recolectar Datos Primarios.....	22
CAPÍTULO 3: PASO A PASO PARA RECOLECTAR DATOS	23
3.1. Identificación de necesidades de información estadística	23
3.2. Diseñar el plan de recolección	26
3.3. Apertura focalizada y gradual	30
CAPÍTULO 4: INDICADORES Y CÓMO MEDIRLOS	37
4.1 Indicadores de la Cadena de Resultados	37
4.2 Indicadores de Gestión	38
CAPÍTULO 5: ¿QUIÉN HACE QUÉ EN LA RECOLECCIÓN?	39
CAPÍTULO 6 DATOS ABIERTOS Y PUBLICACIÓN	43
7.1 ¿Qué son los Datos Abiertos?.....	43
7.2 Características de los Datos Abiertos.....	43
7.3 El Observatorio Regional de Junín	43
7.4 Proceso de Publicación de Datos.....	44
CAPÍTULO 9: DISPOSICIONES FINALES	45
9.1 Disposiciones Complementarias	45
9.2 Disposiciones Transitorias.....	45

SIGLAS

Sigla / Acrónimo	Descripción
CEPLAN	Centro Nacional de Planeamiento Estratégico
DNI	Documento Nacional de Identidad
GRJ	Gobierno Regional de Junín
GPS	Sistema de Posicionamiento Global
SU	Sistema
INEI	Instituto Nacional de Estadística e Informática
ONG	Organizaciones No Gubernamentales
ONGEI	Oficina Nacional de Gobierno Electrónico e Informática
OREJ	Observatorio Regional de Junín
PCM	Presidencia del Consejo de Ministros
PDRC	Plan de Desarrollo Regional Concertado
PEI	Plan Estratégico Institucional
POI	Plan Operativo Institucional
SGCTP	Subgerencia de Cooperación Técnica y Planeamiento
SIAGIE	Sistema de Información de Apoyo a la Gestión de la Institución Educativa
SIS	Sistema de Información de Salud
SIAF	Sistema Integrado de Administración Financiera
SIGA	Sistema Integrado de Gestión Administrativa
SIER	Sistema Integrado de Estadística Regional
UGEL	Unidad de Gestión Educativa Local
UE	Unidad Ejecutora

PRESENTACIÓN

El Gobierno Regional de Junín, en el marco de su compromiso con la modernización de la gestión pública y la implementación de la Gobernanza Digital, presenta la "Guía Práctica para la Recolección de Datos en el Gobierno Regional de Junín".

En un entorno donde la información constituye un activo estratégico fundamental, la capacidad de recolectar, procesar y analizar datos con calidad se vuelve esencial para la toma de decisiones basadas en evidencia. Esta guía ha sido elaborada con el propósito de fortalecer la gestión integral y segura de datos institucionales, garantizando la transparencia, la interoperabilidad y el acceso a la información pública.

El documento está dirigido a todos los servidores del GORE Junín y ofrece un enfoque práctico y amigable, traduciendo conceptos técnicos complejos en instrucciones claras y sencillas. A través de sus capítulos, el lector encontrará herramientas para identificar necesidades de información, aplicar metodologías de recolección adecuadas, distinguir entre tipos de indicadores y cumplir con los estándares de los Datos Abiertos.

Al estandarizar los procesos de recolección mediante formatos oficiales y protocolos de calidad, buscamos no solo cumplir con el marco normativo vigente, sino también transformar los datos en conocimiento útil que impulse el desarrollo regional y mejore los servicios brindados a la ciudadanía.

Zocimo Cárdenas Muje
Gobernador Regional

CAPÍTULO 1:

ASPECTOS GENERALES

1. Objetivo

Fortalecer la gobernanza digital del Gobierno Regional de Junín mediante la gestión integral y segura de datos, orientada a la toma de decisiones basadas en evidencia, la transparencia, la interoperabilidad institucional y la innovación pública, con un enfoque de valor público y protección de datos personales.

2. Alcance

La presente directiva es de cumplimiento obligatorio para todas las unidades orgánicas, órganos desconcentrados, entes adscritos, proyectos especiales que operen datos del Gobierno Regional de Junín y proveedores de sistemas que administren información institucional. Su aplicación comprende tanto las fuentes primarias como los registros administrativos utilizados en la producción, gestión y uso de información estadística.

3. Conceptos

3.1. Fases del proceso estadístico

- 3.1.1. Detección y análisis de necesidades:** Etapa inicial del ciclo de recolección de datos que, mediante revisión de fuentes secundarias y consulta a actores locales, determina qué información falta o es insuficiente para medir brechas sociales, económicas o ambientales, y establece así los temas, cobertura y nivel de desagregación que deben incluirse en los instrumentos de captura regional.
- 3.1.2. Diseño:** Fase en la que se definen de forma anticipada los elementos técnicos y operativos que regirán la captura de información: objetivos de la recolección, universo o muestra, variables e indicadores, instrumentos (formularios, plantillas, apps), flujo de trabajo, roles del equipo, cronograma, presupuesto y protocolos de control de calidad, garantizando que los datos que se obtendrán sean útiles, confiables y compatibles con los estándares regionales y nacionales.
- 3.1.3. Construcción:** Fase del proceso estadístico en la que se elaboran o desarrollan y prueban los mecanismos, los instrumentos, las herramientas, así como los procesos o actividades, siguiendo las especificaciones del diseño, hasta el punto en que están listos para la puesta en funcionamiento.

- 3.1.4. Recolección o acopio:** Fase del proceso estadístico en la que se ejecutan todas las acciones planeadas, diseñadas y construidas en las fases anteriores con el fin de obtener los datos que permitirán generar la información estadística que satisfaga las necesidades identificadas.
- 3.1.5. Procesamiento:** Fase del proceso estadístico en la que se consolidan, integran, procesan y depuran los datos, de acuerdo con lo establecido en el diseño.
- 3.1.6. Evaluación:** Revisión sistemática y en tiempo real de la calidad, cobertura y pertinencia de la información que se está capturando, con el fin de detectar errores, ajustar instrumentos, corregir rutas de campo y garantizar que los datos finales cumplan los estándares establecidos en los lineamientos regionales antes de su almacenamiento y uso.
- 3.1.7. Difusión:** Fase del proceso estadístico en la que se pone a disposición de los usuarios la información estadística, a través de los medios de divulgación establecidos.

3.2. Fases del proceso estadístico

- 3.2.1. Datos estructurados:** información organizada bajo un esquema predefinido: cada atributo ocupa un campo fijo, con tipo de dato, longitud y formato establecidos. Permiten almacenamiento tabular (filas-columnas), consultas rápidas, validación automática y agregación sencilla. Ejemplos: registros de población en una hoja de cálculo, formularios digitales con listas cerradas, bases de datos de colegios o de obras con códigos únicos.
- 3.2.2. Datos no estructurados:** información que carece de esquema previo o cuyo formato es variable y no se ajusta a filas y columnas. Requieren procesamiento adicional (transcripción, codificación, análisis de texto, audio o imagen) para convertirse en datos estructurados. Ejemplos: actas de reuniones escaneadas, grabaciones de entrevistas, fotografías de obras, redes sociales, notas de campo o dictámenes escritos en lenguaje natural.
- 3.2.3. Repositorio de datos:** Almacén digital, interoperable y de acceso controlado, donde el Gobierno Regional conserva, cataloga y pone en valor los conjuntos de datos que genera o recibe en el ejercicio de sus competencias, cumpliendo los principios de seguridad, calidad, reutilización y apertura establecidos en la Ley de Gobierno Digital y su reglamento.
- 3.2.4. Registro administrativo:** conjunto de datos que contiene la información recogida y conservada por las unidades organizacionales en el cumplimiento de sus funciones o

competencias misionales u objetos sociales. Este puede convertirse en una base de datos que se genera y actualiza de forma continua durante la gestión cotidiana de sus servicios (p. ej., matrícula escolar, fichas de salud, licencias de obra, padrón de productores). No requiere diseño estadístico de muestreo, pero sí validación periódica para convertirse en fuente confiable de indicadores regionales.

3.2.5. Anonimización de microdatos: proceso técnico que consiste en transformar los datos individuales de las unidades de observación, de tal modo que no sea posible identificar sujetos o características individuales de la fuente de información, preservando así las propiedades estadísticas en los resultados.

3.2.6. Consistencia: proceso mediante el cual se analiza la relación lógica y numérica entre dos o más variables, con el fin de garantizar la coherencia de la información estadística.

3.3. Componentes estadísticos

3.3.1. Fuente primaria: Es aquella que obtiene los datos directamente de las unidades estadísticas empleando instrumentos propios.

3.3.2. Unidad de observación: Elemento o conjunto de elementos sobre los que se hace la medición de las diferentes variables en una operación estadística.

3.3.3. Unidad de análisis: Elemento de estudio sobre el que se presentan los resultados o las conclusiones de la operación estadística.

3.3.4. Población objetivo: Conjunto de elementos de los que se desea obtener los datos y sobre los que se presentan conclusiones o resultados; está compuesto por unidades que comparten alguna característica, tienen una localización geográfica y un tiempo como periodo de referencia.

3.4. Instrumentos institucionales

3.4.1. Observatorio Regional: Instrumento de gestión y difusión de datos abiertos. Es una plataforma que permitirá realizar el seguimiento de los planes territoriales, el centro de referencia de información, la disponibilidad de información on-line y gratuita, así como el desarrollo de análisis e investigación.

3.5. Estandarización de formatos: Todas las unidades organizacionales deberán usar los formatos y plantillas oficiales aprobados por la Subgerencia de Cooperación Técnica y Planeamiento del Gobierno Regional de Junín (en adelante, GORE Junín) para el acopio y procesamiento de los datos que producen.

- 3.6. Periodicidad:** La recolección de datos se realizará conforme a la periodicidad establecida en los instrumentos de planificación estadística institucional (diaria, semanal, mensual, trimestral o anual), de acuerdo con la naturaleza de la información y la función de cada unidad organizacional.
- 3.7. Validez de la información:** Se considerarán oficiales únicamente los datos debidamente validados por el responsable designado o el planificador de la unidad orgánica del GORE Junín, conforme a los procedimientos de control de calidad y trazabilidad definidos por la Subgerencia de Cooperación Técnica y Planeamiento.
- 3.8. Compatibilidad técnica:** Los registros y bases de datos deberán garantizar su compatibilidad técnica, estructural y semántica con los sistemas institucionales y con el Sistema Integrado de Estadística Regional (SIER), asegurando la interoperabilidad en el marco del Decreto Legislativo N.º 1412 – Ley de Gobierno Digital.
- 3.9. Confidencialidad y protección de datos personales:** Los registros y bases de datos deberán garantizar su compatibilidad técnica, estructural y semántica con los sistemas institucionales y con el Sistema Integrado de Estadística Regional (SIER), asegurando la interoperabilidad en el marco del Decreto Legislativo N.º 1412 – Ley de Gobierno Digital.
- 3.10. Tipos de indicadores y datos generados por el GORE Junín.** Los indicadores de información estratégica y estructural, se encuentran plasmados en el siguiente gráfico:



*Gráfico 1*Articulación de políticas y planes en la cadena de resultados

Nota. Adaptado a la Directiva N° 001-2024-CEPLAN/PCD (Pag. 28). Recuperado de: <https://www.gob.pe/institucion/ceplan/normas-legales/5373394-001-2024-ceplan-pcd>

Los planes territoriales e institucionales se articulan en una única cadena de resultados mediante indicadores alineados y jerarquizados: el Plan Operativo Institucional contiene indicadores de insumos y actividades; el Plan Estratégico Institucional, indicadores de productos y resultados iniciales; y el Plan de Desarrollo Regional Concertado, indicadores de resultados intermedios y finales. Sobre los planes específicos y las políticas

regionales, obedecen a su metodología de diseño que bien podría ir desde indicadores de actividades hasta resultados. Todos los indicadores deben responder a los criterios de verificabilidad, mensurabilidad y correspondencia con la cadena de resultados exigidos por las guías del CEPLAN.

3.10.1. Indicadores de información estratégica.

Son todas aquellas operaciones estadísticas que tengan impacto a nivel distrital, provincial o sectorial al cual pertenece la Unidad Ejecutora, que sean útiles y utilizadas para el diseño, formulación, seguimiento y evaluación de políticas públicas y, por ende, trascienden el ámbito interno de la Entidad. Se encuentran alineados al Plan Estratégico Institucional y al Plan Operativo Institucional.

3.10.1.1. Indicador de insumo y proceso

Los indicadores de estos indicadores abordan los dos primeros componentes de la cadena de resultados —insumos y actividades—. Los indicadores de insumos miden la cantidad de factores de producción (físicos, humanos, financieros, entre otros) que participan o son transformados en el proceso de generación de valor; mientras que los de proceso cuantifican el conjunto de operaciones o actividades necesarias para la producción de bienes o servicios.

- **Características:**

- ✓ Se centran en los medios o inputs necesarios para ejecutar una acción.
- ✓ Son cuantitativos (presupuesto asignado, número de personal, etc.) o cualitativos (calidad de los materiales).
- ✓ Son usados para monitorear la fase inicial de la cadena de valor de una política o proyecto.

- **Ejemplo:**

- ✓ Monto ejecutado en proyectos de infraestructura educativa (S/).
- ✓ N° de profesionales de salud contratados.
- ✓ Monto destinado al mantenimiento de carreteras (S/).
- ✓ N° de equipos médicos adquiridos.
- ✓ N° de kits escolares distribuidos.

3.10.1.2. Indicadores de Producto:

Los indicadores de producto miden la provisión de un bien o servicio a la población objetivo por parte de la entidad. La medición de este indicador implica que la población objetivo reciba el bien o servicio cumpliendo ciertos atributos (estándares) previamente definidos.

- **Características:**

- ✓ Se centran en los entregables o productos generados.
- ✓ Pueden ser cuantitativos (número de escuelas construidas) o cualitativos (calidad de los servicios entregados).

- ✓ Sirven como puente entre los insumos utilizados y los resultados esperados.

- **Ejemplo:**

- ✓ Número de docentes capacitados en metodologías activas.
- ✓ Número de campañas de vacunación realizadas.
- ✓ Número de productores capacitados en buenas prácticas agrícolas.
- ✓ Kilómetros de carretera rehabilitados o mantenidos.
- ✓ Hectáreas reforestadas con especies nativas.
- ✓ Número de ferias agropecuarias realizadas

3.10.2. Indicadores estructurales

Los indicadores estructurales son indicadores estratégicos, pero su alcance es independiente del período de gobierno. Se encuentran ligados a resultados de largo plazo. Ligados al Plan de Desarrollo Regional Concertado.

3.10.2.1. Indicadores de Resultado:

Los indicadores de resultados miden los efectos o cambios que se han generado en la población o el territorio como consecuencia de los productos entregados. Reflejan el impacto de las acciones del GORE Junín en el cumplimiento de los objetivos estratégicos, como la mejora de las condiciones de vida o el desarrollo regional.

3.10.2.1.1. Indicador de resultado inicial

Un indicador será una medida de resultado inicial si permite cuantificar los cambios en las condiciones de vida o estado de la población objetivo en el corto plazo como resultado de la intervención de una entidad. Esta ejerce influencia directa sobre el indicador, por tanto, se considera que los resultados iniciales son resultados institucionales.

- **Características:**

- ✓ Enfocados en cambios observados en el comportamiento, el acceso o la cobertura inicial.
- ✓ La temporalidad, a corto plazo.
- ✓ Sirve como base para evaluar el progreso hacia resultados intermedios y finales, plasmados en el Plan de Desarrollo Regional Concertado.

- **Ejemplo:**

- ✓ Porcentaje de niños menores de 3 años que reciben suplementos de hierro en centros de salud de Junín.
- ✓ Porcentaje de familias con acceso a agua potable.
- ✓ Número de productores capacitados en buenas prácticas agrícolas.

- ✓ Porcentaje de vías regionales en buen estado de transitabilidad.
- ✓ Porcentaje de viviendas rurales con energía eléctrica.

3.10.2.1.2. Indicador de resultado intermedio

Un indicador será una medida de resultado intermedio si permite cuantificar los cambios en las condiciones de vida o estado de la población objetivo en el mediano plazo. Una entidad solo tendrá influencia indirecta sobre este tipo de resultados; por tanto, para su consecución se requiere la intervención de varias entidades y otros actores de la sociedad.

- **Ejemplo:**

- ✓ Porcentaje de estudiantes que alcanzan nivel satisfactorio en comprensión lectora.
- ✓ Porcentaje de población que accede a servicios de salud preventivos.
- ✓ Porcentaje de mujeres que realizan control prenatal completo.
- ✓ Porcentaje de hogares con ingresos por encima de la línea de pobreza.
- ✓ Porcentaje de agricultores que aplican tecnologías sostenibles.

3.10.2.1.3. Indicador de resultado final

Un indicador será una medida de resultado final si permite cuantificar los cambios de las condiciones de vida de la población objetivo a largo plazo y requiere de la participación de varias entidades para su consecución. Un resultado final tendrá como determinantes principales un conjunto de resultados intermedios, por tanto, su nivel de complejidad es alto.

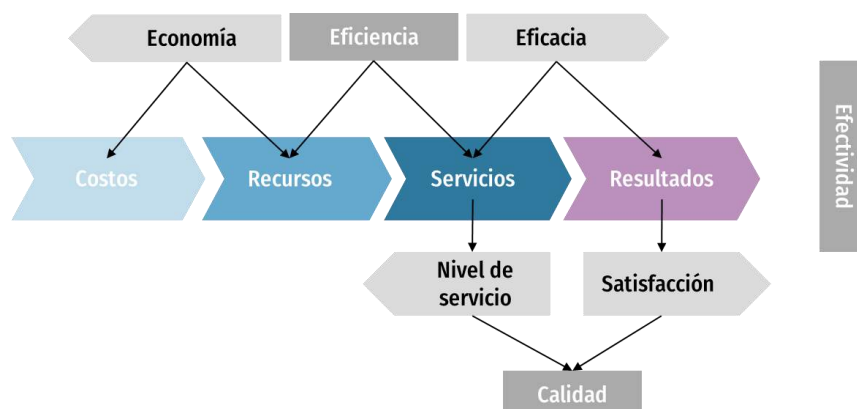
- **Ejemplo:**

- ✓ Tasa de desnutrición crónica infantil.
- ✓ Tasa de mortalidad materna.
- ✓ Tasa de pobreza total en la región.
- ✓ Esperanza de vida al nacer.
- ✓ Tasa de empleo digno.
- ✓ Porcentaje de población con educación superior completa.
- ✓ Nivel de competitividad regional.

3.10.3. Indicadores de gestión

Un indicador de gestión es una medida cuantitativa o cualitativa que muestra el grado de avance, eficiencia, eficacia, economía o calidad con que el GORE Junín ejecuta sus actividades, utiliza sus recursos y alcanza sus objetivos. Estos indicadores evidencian qué tan bien está funcionando el GORE Junín en relación con lo que se planificó.

Se determina de la siguiente manera:



3.10.3.1. Indicadores de eficacia

Cuantifican el grado de cumplimiento de los objetivos o resultados de la institución, programa o proyecto sin considerar los recursos utilizados. Las medidas de eficacia serán útiles en la medida en que los objetivos y/o resultados de la institución o programa se encuentren claramente determinados.

- **Ejemplo:**

- ✓ Porcentaje de estudiantes de secundaria que aprueban el examen de competencias básicas.
- ✓ Porcentaje de kilómetros de carreteras rurales rehabilitadas según lo planificado en Junín.
- ✓ Porcentaje de reducción de la tasa de anemia infantil.
- ✓ Porcentaje de reducción de la pobreza extrema en zonas rurales.
- ✓ Porcentaje de metas institucionales alcanzadas respecto al PEI.

3.10.3.2. Indicadores de eficiencia

Describe la relación entre dos magnitudes, la producción de un bien o servicio y los insumos (financieros, humanos o bienes) utilizados para su generación. Este tipo de indicador está referido a la producción de bienes o servicios, utilizando el mínimo de recursos disponibles.

- **Ejemplo:**

- ✓ Costo por kilómetro de carretera construido
- ✓ Costo de un servicio provisto por usuario
- ✓ Cantidad de inspecciones realizadas en relación al total de inspectores.
- ✓ Costo promedio por estudiante atendido.
- ✓ Costo promedio por paciente atendido.

3.10.3.3. Indicadores de calidad

Este indicador se encuentra orientado exclusivamente a medir la capacidad de la institución o programa para responder ante las necesidades de sus clientes, usuarios o población objetivo. Está referido a las características y/o atributos de los bienes y servicios entregados, como la oportunidad, la accesibilidad, la continuidad, la cortesía en la atención, entre otros.

- **Ejemplo:**

- ✓ Tiempo de espera promedio en oficinas de atención al público.
- ✓ Nivel de satisfacción de los padres de familia con el servicio educativo.
- ✓ Usuarios satisfechos en relación al total de usuarios
- ✓ Tiempo de respuesta promedio a cartas de reclamos.
- ✓ % de usuarios que perciben mejoras en los servicios públicos locales.

3.10.3.4. Indicadores de economía

Este indicador se relaciona con la capacidad de una institución, programa o intervención, para administrar sus recursos financieros, generar ingresos propios para financiar sus actividades, recuperar préstamos.

- **Ejemplo:**

- ✓ Presupuesto ejecutado respecto al presupuesto programado.
- ✓ Porcentaje de recuperación de préstamos.
- ✓ Porcentaje de recursos generados en relación al gasto total para la producción del bien o servicio.
- ✓ Porcentaje de ingresos propios respecto al total de ingresos (%)
- ✓ Porcentaje de recursos ordinarios sobre el total de ingresos (%)

3.10.4. Datos abiertos

Los datos abiertos son aquellos datos primarios o sin procesar que las entidades públicas ponen a disposición de cualquier ciudadano en formatos abiertos, interoperables y de libre acceso, sin restricciones de uso. Su finalidad es promover la transparencia, la participación ciudadana y la innovación pública, conforme al Decreto Supremo N.º 016-2017-PCM y al Decreto Legislativo N.º 1412 – Ley de Gobierno Digital

3.10.5. Datos geoespaciales

Un dato espacial es aquel que tiene asociada una referencia geográfica, ya sea directa mediante coordenadas o indirecta mediante un código postal, de modo que se puede localizar exactamente dónde ocurre en un mapa. Los datos espaciales aportan información sobre hechos vinculados a objetos que tienen una extensión en el espacio, entendiendo

por extensión el área que cubre un objeto y que se caracteriza por tamaño, posición y forma.

4. Participantes del proceso

4.1. La Subgerencia de Cooperación Técnica y Planeamiento tiene las siguientes responsabilidades:

- Complementa con lineamientos la presente directiva, los cuales serán aprobados mediante resolución de la Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial.
- Formula, aprueba y difunde los lineamientos, formatos y cronogramas oficiales para la recolección de datos.
- Coordina con todas las unidades para asegurar la estandarización de instrumentos y procedimientos.
- Consolida, valida y publica la información dentro de los plazos establecidos.
- Supervisa y evalúa el cumplimiento de los procesos y propone mejoras continuas.
- Desarrolla y mantiene el Repositorio Estadístico Regional como parte del Observatorio Regional de Junín.

4.2. Las Direcciones Regionales y Gerencias:

- Implementa los lineamientos y formatos oficiales en su ámbito de competencia.
- Designa un responsable de información que actúe como enlace.
- Organiza y supervisa las actividades de recolección de datos de sus dependencias.
- Garantiza la prevalidación técnica y de coherencia de los datos a través de sus planificadores.
- Garantiza la disponibilidad de recursos humanos y materiales para la ejecución del levantamiento de información.
- Valida internamente la información antes de su remisión a la Subgerencia de Cooperación Técnica y Planeamiento.

4.3. Las unidades ejecutoras/unidades organizacionales tienen las siguientes responsabilidades:

- Planifica y ejecuta el levantamiento de datos según los cronogramas y metodologías.
- Designa un equipo de trabajo responsable de la captura, el registro y procesamiento.
- Mantiene un registro ordenado y documentado de todas las fuentes y evidencias.
- Asegura que los datos recolectados sean completos, consistentes y veraces.

- Realiza controles de calidad internos antes de enviar la información.

4.4. La Oficina de Planeamiento de cada unidad organizacional:

- Integra la información recolectada del centro de producción de datos.
- Verifica la coherencia entre los datos y las metas institucionales.
- Coordina con la Subgerencia de Cooperación Técnica y Planeamiento para resolver observaciones o inconsistencias.

4.5. La Oficina Regional de Desarrollo Institucional y Tecnologías de Información, es responsable de:

- Provee el soporte técnico y las herramientas digitales para el registro y la transmisión segura de datos.
- Mantiene operativos los sistemas y aplicaciones que facilitan la captura de información.
- Capacita a los usuarios en el manejo de las plataformas digitales.
- Publica la información en la plataforma de datos abiertos, como parte del gobierno digital.

4.6. Personal de campo y/o encuestadores:

- Cumple estrictamente con los protocolos y formatos aprobados para la recolección.
- Garantiza la confidencialidad y protección de los datos personales.
- Reporta cualquier incidencia que pueda afectar la calidad de los datos.
- Reporta semanalmente el avance a la Subgerencia de Cooperación Técnica y Planeamiento.

CAPÍTULO 1:

¿QUÉ ES LA RECOLECCIÓN DE DATOS?

1.1. Definición:

La recolección de datos es como hacer una encuesta o tomar notas sobre algo que observamos. Imagina que quieres saber cuántas personas viven en tu barrio, qué servicios de salud usan o qué problemas tienen. Para averiguarlo, necesitas preguntar, observar y anotar. Ese proceso de preguntar, observar y registrar la información se llama recolección de datos.

En el Gobierno Regional de Junín, recolectamos datos todos los días: cuando registramos a los beneficiarios de un programa, cuando contamos cuántas personas atendimos, cuando medimos el avance de una obra o cuando hacemos una encuesta de satisfacción. Todos estos son ejemplos de recolección de datos.

1.2. ¿Por qué es importante recolectar datos bien?

Recolectar datos de forma correcta es muy importante por varias razones. Aquí te explicamos las principales:

RAZÓN	EXPLICACIÓN SENCILLA
Para tomar buenas decisiones	Si los datos son correctos, las autoridades pueden decidir dónde construir un hospital, qué lugares necesitan más apoyo, o cuánto dinero destinar a cada proyecto. Es como tener un mapa correcto para no perderse.
Para conocer la realidad	Los datos nos dicen cómo está nuestra región: cuántos niños van a la escuela, qué enfermedades son más comunes, qué tantas carreteras están en buen estado. Sin datos, trabajaríamos a ciegas.
Para rendir cuentas	Los ciudadanos tienen derecho a saber qué hace el gobierno. Con buenos datos, podemos mostrar claramente qué hemos hecho y qué resultados hemos obtenido.
Para cumplir la ley	Existen normas que nos obligan a tener registros correctos y a proteger la información personal de los ciudadanos. Recolectar bien nos ayuda a cumplir estas normas.
Para no repetir trabajo	Si recolectamos bien los datos desde el inicio, no tendremos que volver a preguntar lo mismo después. Ahorramos tiempo, dinero y esfuerzo.

1.3. Las 6 Fases del Proceso de Recolección

La recolección de datos no es algo que se haga de cualquier manera. Sigue un proceso ordenado que tiene varias fases. Si seguimos las fases correctamente, tendremos información de calidad.

LAS 7 FASES DE LA RECOLECCIÓN DE DATOS

FASE 1: IDENTIFICACIÓN DE LA INFORMACIÓN

Primero debemos preguntarnos: ¿Qué información necesitamos? ¿Qué nos falta saber? Es como hacer una lista de mercado antes de ir a comprar.

FASE 2: ANÁLISIS DE LA INFORMACIÓN

Aquí decidimos cómo vamos a obtener los datos: qué preguntas haremos, a quién se las haremos, con qué herramientas trabajaremos.

FASE 3: PRIORIZACIÓN DE LA INFORMACIÓN

Preparamos los formularios, cuestionarios o fichas que usaremos. También probamos que funcionen bien antes de usarlos.

FASE 4: PROCESAR LA INFORMACIÓN

Organizamos los datos, los pasamos a archivos digitales, corregimos errores y los preparamos para el análisis.

FASE 5: DOCUMENTACIÓN DE LOS DATOS

Revisamos que los datos estén bien, que no haya errores, que estén completos y que tengan sentido.

FASE 6: DIFUNDIR LOS RESULTADOS

Finalmente, publicamos los datos para que todos puedan verlos y usarlos. Es el momento de compartir nuestros hallazgos.

CAPÍTULO 2:

TIPOS DE DATOS Y CÓMO RECOLECTARLOS

2.1. Datos Estructurados vs No Estructurados

No todos los datos son iguales. Algunos son fáciles de organizar en tablas y otros son más complicados. Es importante que conozcas la diferencia para saber cómo tratar cada tipo de dato.

TIPO	¿QUÉ ES?	EJEMPLOS
DATOS ESTRUCTURADOS	Son datos que se pueden organizar en filas y columnas, como una tabla de Excel. Tienen un formato fijo y ordenado.	• Lista de estudiantes con nombre, edad y grado • Registro de pacientes con DNI, diagnóstico y fecha • Padrón de productores con hectáreas y cultivos
DATOS NO ESTRUCTURADOS	Son datos que no se pueden poner fácilmente en una tabla. Vienen en formatos libres como texto, audio o imágenes.	• Actas de reuniones escaneadas • Grabaciones de entrevistas • Fotografías de obras • Notas de campo escritas a mano

2.2. Fuentes de Datos: Primarias y Secundarias

Cuando necesitamos datos, podemos obtenerlos de dos formas: crearlos nosotros mismos (fuente primaria) o usar datos que ya existen (fuente secundaria). Cada opción tiene sus ventajas y desventajas.

TIPO DE FUENTE	¿QUÉ ES?	¿CUÁNDO USARLA?
FUENTE PRIMARIA	Datos que obtenemos directamente de la realidad. Salimos a preguntar, observar o medir. Son datos frescos, hechos a la medida.	Cuando no existe información previa, cuando necesitamos datos actualizados o muy específicos de nuestra región.
FUENTE SECUNDARIA	Datos que ya fueron recolectados por otras instituciones como el INEI, ministerios u otras oficinas. Solo los consultamos.	Cuando la información ya existe y es confiable, cuando tenemos poco tiempo o presupuesto, o para

complementar nuestros datos.

2.3. Métodos para Recolectar Datos Primarios

Cuando necesitamos obtener datos directamente de la fuente, podemos usar diferentes métodos. Aquí te explicamos los principales y cuándo usar cada uno:

MÉTODO	¿EN QUÉ CONSISTE?	EJEMPLO DE USO
CENSO	Preguntar a TODAS las personas o elementos del grupo que estudiamos. Es completo, pero puede ser costoso.	Inventario de todas las escuelas de una provincia. Registro de todos los médicos de la red de salud.
ENCUESTA	Preguntar a una PARTE representativa del grupo. Es más económico y nos da resultados confiables si se hace bien.	Medir el nivel de satisfacción de usuarios de servicios de salud. Conocer las necesidades de productores agrícolas.
OBSERVACIÓN DIRECTA	Ir al lugar y anotar lo que vemos. No preguntamos, solo observamos y registramos hechos.	Verificar el estado de una carretera. Observar el flujo de personas en un mercado.
MEDICIÓN FÍSICA	Usar instrumentos para medir algo: GPS, balanzas, termómetros, cintas métricas, sensores.	Medir el área de un terreno. Pesar la producción agrícola. Medir la calidad del agua.
REVISIÓN DE REGISTROS	Sacar datos de documentos que ya existen: actas, libros, archivos, sistemas informáticos.	Consultar el registro de atenciones médicas. Revisar los partes diarios de una obra.

CAPÍTULO 3:

PASO A PASO PARA RECOLECTAR DATOS

Ahora que ya conoces los conceptos básicos, te explicamos el proceso completo para recolectar datos correctamente. Sigue estos pasos en orden y tendrás información de calidad.

3.1. Identificación de necesidades de información estadística

La Subgerencia de Cooperación Técnica y Planeamiento, a través de la Unidad de Estadística y Gestión del Conocimiento, recepciona e identifica las necesidades de información de los usuarios de la información estadística externos y a solicitud para realizar el seguimiento y evaluación de los planes institucionales y territoriales del GORE Junín; con el fin de priorizar y garantizar que la producción de la información estadística se encuentre alineada con las necesidades previamente identificadas (Ver Formato 1).

Antes de salir a recolectar datos, debemos estar seguros de qué es lo que necesitamos. No tiene sentido preguntar cosas que no vamos a usar. Para detectar las necesidades de información, hazte estas preguntas:

- ¿Qué decisión importante necesitamos tomar? (Ejemplo: ¿Dónde construir un nuevo centro de salud?)
- ¿Qué datos nos faltan para tomar esa decisión? (Ejemplo: No sabemos cuántas personas viven en cada distrito)
- ¿Ya existe esa información en alguna parte? (Revisa el INEI, ministerios, o archivos de la institución)
- ¿Quiénes necesitan esa información? (Otros gerentes, ciudadanos, entidades cooperantes)

3.1.1. Caracterización de usuarios.

Se realizará un mapeo de usuarios de información estadística regional interesados en la producción de datos. Para ello, se usará un cuestionario para el acopio de información. La Oficina de Comunicaciones, en coordinación con la Oficina Regional de Desarrollo Institucional y Tecnologías de Información, realiza las publicaciones por los medios virtuales pertinentes; su análisis y procesamiento serán realizados por la Subgerencia de Cooperación Técnica y Planeamiento a través de la Unidad de Estadística y Gestión del Conocimiento; este proceso se clasificará por tipo, nivel, frecuencia de uso y temas de interés:

Tipo de usuario	Ejemplos	Temas de interés
Interno	Gerencias, Direcciones, Unidades de Gestión, Redes de Salud	Indicadores de gestión institucional
Gobiernos Locales	Municipalidades Provinciales / Distritales	Obras, metas presupuestales,

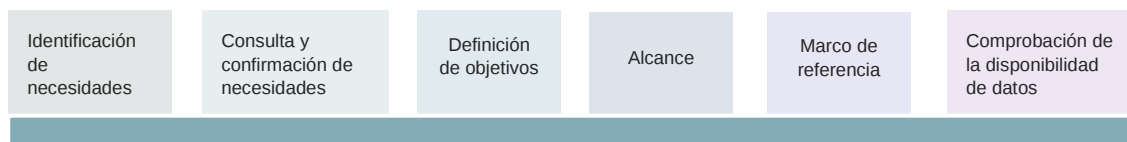
		estadísticas territoriales
Academia	Universidades, institutos y centros de investigación.	Desarrollo regional, pobreza, medio ambiente.
Sociedad civil	Observatorios ciudadanos o sectoriales, ONG y medios.	Transparencia, derechos humanos, servicios
Sector privado	Cámaras de comercio, gremios empresariales y asociaciones productivas	Inversión, empleo, productividad, competitividad.
Organismos desconcentrados y cooperantes	Banca bilateral o multilateral, entes cooperantes	Inversión territorial.

La presente clasificación es referencial y podrá ampliarse o actualizarse conforme a las necesidades emergentes de información identificadas en los procesos de planificación y monitoreo regional.

3.1.2. Necesidades de información.

Para identificar las necesidades de información estadística se considera por igual a los que están dentro del GORE Junín —sus planes, políticas y equipos de trabajo— y a los de afuera: la gente de a pie, la academia, las empresas y otros aliados. Todo parte de las solicitudes que llegan a la Subgerencia de Cooperación Técnica y Planeamiento, que conversa con quienes producen los datos; también revisamos cómo los indicadores encajan con nuestros planes (PEI, POI, PDRC), con las políticas públicas y con lo que piden los ciudadanos a través de consultas, encuestas, audiencias o redes oficiales. Al final, elegimos lo más urgente según lo que más nos importa como institución, si los datos ya existen y son confiables, si van con nuestros objetivos y cuánta gente los está pidiendo.

El proceso de identificación de necesidades se desarrolla según las etapas que se muestran en el siguiente flujo:



El proceso nace de la identificación de las necesidades internas y externas de datos por la Unidad de Estadística y Gestión de Conocimientos de la SGCTP, teniendo una lista de necesidades se analiza a fin de realizar la consulta a las áreas generadoras de datos y revisión del repositorio estadístico concluyendo en la confirmación de necesidades, para ello se pretende determinar los objetivos claros con la

producción de datos, asimismo, se determina los alcances de su acopio, y se desarrolla un marco referencia, llegando a la comprobación de disponibilidad de datos.

Para identificar las necesidades de información estadística del GORE Junín, se aplicarán diversos mecanismos de consulta dirigidos a los usuarios internos y externos del sistema estadístico regional. Estos mecanismos permitirán recoger demandas, evaluar brechas de información y establecer prioridades de producción estadística.

La Subgerencia de Cooperación Técnica y Planeamiento será responsable de coordinar la aplicación de los instrumentos, consolidar los resultados y elaborar el informe técnico de priorización.

A continuación, se detallan los mecanismos a emplear en el proceso:

Mecanismo	Público objetivo	Herramienta	Frecuencia
Encuesta de necesidades	Usuarios externos	Formulario digital	Anual
Talleres sectoriales	Internos + Externos	Presencial o virtual	Trimestral
Reunión de coordinación	Internos	Acta	Mensual
Solicitudes	Todos	Formulario Web o escrito	Continuo
Observatorio ciudadano	Sociedad civil	Lluvia de ideas	Semestral

Una vez identificadas las necesidades de información, se deberá elaborar un documento que contenga el análisis y la priorización de dichas necesidades, considerando su área temática de estudio, la relación con la misionalidad de la entidad, el alcance previsto de la operación estadística, el marco normativo aplicable y la disponibilidad de recursos para su ejecución.

El resultado de este análisis permitirá orientar la programación de actividades estadísticas y fortalecer la alineación entre la oferta de datos institucionales y las necesidades reales de los usuarios.

3.1.3. Identificar la información disponible y las brechas.

Este proceso tiene por finalidad determinar qué información estadística existe actualmente en las unidades orgánicas del GORE Junín (ver formato 6), en qué formato se encuentra, su nivel de calidad y qué datos

faltan para atender los objetivos de planificación, inversión o difusión institucional.:

- **Inventario de fuentes internas.** Verifica, mediante el uso de un checklist de calidad (ver Formato 4), la información existente en los registros administrativos, informes y bases de datos de las unidades organizacionales del GORE Junín.
 - Registros administrativos (SIAF, SIGA, libros de actas, partes diarias).
 - Reportes operativos (actas de supervisión, informes de control interno).
 - Inventarios o catastros (infraestructura, equipamiento, personal).
 - Sistemas y archivos digitales: bases en Excel y sistemas sectoriales (SIAGIE, HIS, entre otros)
- En caso de encontrar datos no procesados, se deben ejecutar los siguientes pasos: i) tabulación (organizar en tablas), ii) depuración (corregir o eliminar errores), y iii) clasificación (agrupar en categorías homogéneas).

La tabulación consiste en organizar los datos recolectados en tablas o matrices, de forma que puedan visualizarse y analizarse fácilmente; mientras que la depuración es el proceso mediante el cual se verifica, corrige o elimina información inconsistente, duplicada, incompleta o errónea dentro de una base de datos; y la clasificación implica ordenar y agrupar los datos en categorías o tipologías homogéneas, de acuerdo con criterios predefinidos, tales como territorio, población, actividad económica, nivel educativo, entre otros.
- Se usará la lista de verificación para la recolección de datos de fuentes primarias para asegurar la calidad y pertinencia de la información. Considerando mínimamente: la pertinencia (La información recolectada responde al objetivo del levantamiento), confiabilidad (La fuente de datos ha sido validada por la unidad técnica correspondiente), exactitud (Los datos fueron registrados sin errores u omisiones), consistencia (Los datos coinciden con registros administrativos relacionados) y oportunidad (La información corresponde al periodo de análisis establecido).
- Se remitirá un informe a la Subgerencia de Cooperación Técnica y Planeamiento a fin de que pueda ser validado y su posterior producción de datos o indicadores requeridos (Ver Formato 5).

3.2. Diseñar el plan de recolección

Ya sabemos qué necesitamos. Ahora toca planificar cómo lo vamos a conseguir. Un buen plan debe incluir los siguientes elementos:

ELEMENTO DEL PLAN	¿QUÉ DEBES DEFINIR?
Objetivo claro	¿Para qué vamos a recolectar estos datos? ¿Qué queremos lograr con esta información?
Población objetivo	¿A quiénes vamos a estudiar? ¿Todos o solo una muestra? ¿Qué características deben tener?
Variables a medir	¿Qué datos específicos vamos a recolectar? (edad, sexo, ingresos, número de hijos, etc.)
Método de recolección	¿Cómo vamos a obtener los datos? ¿Encuesta, observación, revisión de archivos?
Instrumentos	¿Qué herramientas usaremos? ¿Cuestionario, ficha de campo, formato digital?
Cronograma	¿Cuándo empezamos? ¿Cuándo terminamos? ¿Cuánto tiempo tomará cada actividad?
Responsables	¿Quiénes van a recolectar los datos? ¿Quién va a supervisar? ¿Quién va a procesar?
Presupuesto	¿Cuánto vamos a gastar? ¿En qué? (movilidad, materiales, pagos a encuestadores)

El diseño de recolección y acopio de información comprende la definición de los procedimientos, instrumentos y responsabilidades necesarios para obtener y organizar datos relevantes que sustenten la toma de decisiones institucionales. En esta etapa se determinan las fuentes de información (primarias y secundarias), los métodos de recolección, los criterios de calidad y pertinencia, así como los mecanismos de verificación y control. Asimismo, se establece la aplicación de listas de verificación y formatos estandarizados que garanticen la confiabilidad, consistencia y oportunidad de la información. Para este caso se usan los Anexos 10 y 11. Se seleccionará la técnica de campo más adecuada:

3.2.1. Definir el tipo de fuente primaria y el enfoque de levantamiento.

Se entiende como fuente primaria la información obtenida directamente de la realidad (observación, medición, encuestas, entrevistas) con justificación técnica.

El uso de fuentes primarias se justifica técnicamente por la necesidad de obtener información directa, actualizada y específica que refleje con precisión la realidad del ámbito de intervención del GORE Junín. Este tipo de fuente permite recolectar datos de primera mano, a través de instrumentos diseñados conforme a los objetivos del proceso de

planeamiento, garantizando así la pertinencia, confiabilidad y representatividad de la información.

3.2.1.1. Elección del enfoque:

- **Censo:** Aplicación al 100% de la población objetivo. Se utiliza cuando el universo es reducido o se requiere máxima precisión. Ejemplo:
 - ✓ Inventario de todas las instituciones educativas en el distrito/UGEL/provincia o departamento.
 - ✓ Inventario de los bofedales existentes en el territorio.
- **Encuesta:** Aplicación a una muestra representativa con diseño estadístico (Ver Anexo 7). Se utiliza cuando el universo es muy grande y los recursos son limitados. Ejemplo:
 - ✓ Nivel de satisfacción de la prestación de los servicios que ofrecen los establecimientos de salud.
 - ✓ Escala de valoración del servicio que oferta el GORE Junín.
 - ✓ La Oficina de Planeamiento de las Unidades Ejecutoras o Subgerencia de Cooperación Técnica y Planeamiento determinará la elección muestral, en caso de que no cuenten con dichos instrumentos.

Para determinar el procedimiento, véase el Anexo 12.

- **Registro Administrativo:** Recolección sistemática y continua de datos como parte de las operaciones diarias. Proporciona información actualizada en tiempo real; dentro de las ventajas está la información actualizada y en tiempo real.
 - Registro de atención en un centro de salud.
 - Número de hectáreas de siembra de papa por estación.
 - Toneladas de producción de trucha en el departamento.
 - Exportación de artesanías en soles.

3.2.1.2. Determinar la unidad de análisis y el sujeto de medición

Para determinar la unidad de análisis, se puede determinar:

- 3.2.1.2.1. **Unidad de análisis:** La unidad de análisis corresponde al objeto o entidad sobre la cual se recopilará la información. Puede tratarse de personas, hogares, instituciones, establecimientos, infraestructuras, activos físicos, territorios u otros elementos observables.
- 3.2.1.2.2. **Sujeto de medición:** El sujeto de medición es la unidad de análisis o elemento específico sobre el cual se recolecta información en un proceso de medición, observación o levantamiento de datos. Representa al objeto, individuo, grupo, institución, territorio o fenómeno que se evalúa para generar indicadores y evidencias que sustenten el

planeamiento, seguimiento o evaluación de la gestión pública. Dentro de sus características se encuentran:

- Definición clara: incluir criterios de inclusión y exclusión. Para ello, los estadistas / planificadores o la Unidad de Estadística y Gestión del Conocimiento de la SGCTP podrán desarrollar dichos criterios.
- Delimitación geográfica: departamento, provincia, distrito, centro poblado.
- Delimitación temporal: periodo de referencia (por ejemplo, año, ciclo escolar).
- Pertinencia: alineación con los objetivos que se pretenden alcanzar.

3.2.1.3. Selección de la técnica de recolección.

3.2.1.3.1. Cuestionario estructurado: Es la aplicación de un cuestionario con preguntas cerradas o mixtas a una muestra o población completa. Puede desarrollarse en papel o por medio digital. Dentro de sus limitaciones se encuentra la existencia de sesgo si no hay entrenamiento del encuestador. Para su aplicación deberá ser validada por la Unidad de Estadística y Gestión del Conocimiento de la SGCTP

Dentro de las ventajas, se encuentran:

- ✓ Existe una alta estandarización.
- ✓ Permite validación en línea (rango, saltos, consistencia)
- ✓ El procesamiento es rápido si es digital.
- ✓ Es reutilizable en el tiempo.

3.2.1.3.2. Observación directa: Es el registro sistemático de hechos o condiciones mediante la lista de chequeo o formatos de campo, sin intervención del entrevistado.

Dentro de las ventajas, se encuentran:

- ✓ No depende de la memoria o percepción.
- ✓ Útil para verificar condiciones físicas.
- ✓ Bajo costo si se usa personal técnico institucional.

3.2.1.3.3. Medición física: Es el uso de herramientas o equipos para medir directamente las variables. Se realiza mediante el uso de instrumentos, como GPS, sensores y otros.

3.2.1.3.4. Revisión documental en campo: Consiste en la extracción de datos a partir de registros administrativos, planillas, aplicativos internos, libros. Se aprovecha de la información ya disponible, pero depende de la calidad y actualización de los registros.

3.2.1.4. Diseño del instrumento

3.2.1.4.1. Ficha de recolección: En este proceso debe contener:

- Datos de identificación (fecha, encuestador, código de unidad organizacional). Su diseño depende del tipo de fuentes y técnicas de levantamiento definidas previamente. Para tal caso la Unidad de Estadística y Gestión del conocimiento determinará dichas fichas.
- Cuando existan variables alineadas al indicador (nombre, definición, unidad de medida, periodicidad).
- Campos de observación y validación.

3.2.1.4.2. Manual del encuestador / registrador

- Se contará con un manual estandarizado del encuestador/registrator que explique cada pregunta, cómo registrar y qué hacer ante datos faltantes. Dichos lineamientos serán brindados por la Subgerencia de Cooperación Técnica y Planeamiento.

3.3. Apertura focalizada y gradual

3.3.1. FASE I: Identificación de la información

a) Identificación de la información bajo responsabilidad de las unidades ejecutoras.

Las áreas generadoras de datos elaborarán un listado de la información que producen y/o administran en cumplimiento de sus funciones y lo reportarán al planificador/estadístico de la Unidad Ejecutora.

b) Consolidación del inventario de información institucional

El analista estadístico/planificador consolidará la información en un solo documento para su posterior análisis.

El inventario de información institucional debe ser actualizado permanentemente. Los datos también constituyen un activo de información dentro del GORE Junín, para la toma de decisiones. Para determinar el procedimiento, véanse los anexos **10 y 11**.

3.3.2. FASE II: Análisis de la información

a) Identificación de la información publicable y no publicable, restricciones existentes:

La Subgerencia de Cooperación Técnica evaluará la información para determinar de acuerdo con los criterios jurídicos si existe alguna restricción legal para ser considerada como no publicable, anotando su restricción en el Formato 3. Esto, concordante con el marco legal que regula la confidencialidad y publicación de datos (Ley N.º 27806 – Ley de Transparencia y Acceso a la Información Pública y Ley N.º 29733 – Ley de Protección de Datos Personales).

De la misma forma, el personal de las Áreas generadoras debe identificar las restricciones de naturaleza técnica que pueden limitar la liberación de los datos, anotando su restricción en el Formato.

b) Preparación del inventario de información institucional.

El analista estadístico de la Subgerencia de Cooperación Técnica y Planeamiento consolidará la información resultado del análisis en un solo documento para su posterior priorización (Ver anexos 13 y 14).

Es imprescindible que el Gobierno Regional de Junín, antes de publicar o compartir los datos recolectados, realice un proceso de análisis, clasificación y validación de la información, con el fin de determinar qué datos pueden ser difundidos y cuáles deben mantenerse en reserva. Este procedimiento debe garantizar el cumplimiento de la Ley N.º 29733, Ley de Protección de Datos Personales, y de las normas sobre confidencialidad de la información pública. En los casos que correspondan, la información podrá ser transformada mediante técnicas de anonimización o agregación, que aseguren la protección de la identidad y privacidad de las personas, preservando al mismo tiempo el valor analítico y estadístico de los datos para la gestión institucional.

3.3.3. FASE III: Priorización de la información

En esta fase se pretende definir los conjuntos de datos y el cronograma de la publicación, teniendo en cuenta que se trata de un proceso permanente y gradual. En suma, se tendrá un catálogo de datos a publicar según los criterios de impacto y dificultad.

a) Revisión de los criterios de evaluación.

El equipo de trabajo debe conocer los criterios de evaluación de impacto y dificultad y el procedimiento a seguir.

b) Calificación de la información identificada

El equipo de trabajo reunido efectuará la calificación de la información identificada en el inventario de información institucional.

c) Ordenamiento de los conjuntos de datos que conformarán el catálogo de datos

El equipo de trabajo procederá a ordenar los conjuntos de datos en grupos de corto, mediano y largo plazo.

Cabe resaltar la importancia de iniciar la apertura de datos de forma focalizada y gradual, considerando aquellos conjuntos de datos que produzcan un mayor impacto e impliquen menor dificultad, con ello se asegura evitar controversia y complejidad que podría retrasar el proyecto (Ver anexo 5). Adicionalmente, puede existir variabilidad en cuanto a los criterios de priorización que puedan ser utilizados por las Entidades. Por

ejemplo, en algunos casos se considera la viabilidad técnica e institucional.

3.3.4. FASE IV: Procesar información

Ya terminamos de recolectar. Ahora toca organizar toda la información para que pueda ser analizada. Este proceso incluye varias actividades:

PROCESO DE PROCESAMIENTO DE DATOS

PASO 1: DIGITACIÓN

Si los datos están en papel, hay que pasarlos a la computadora. Usa siempre las plantillas digitales que te proporciona la SGCTP.

PASO 2: TABULACIÓN

Organiza los datos en tablas ordenadas. Cada columna debe tener un encabezado claro y cada fila debe ser un registro completo.

PASO 3: DEPURACIÓN

Revisa y corrige errores. Busca datos duplicados, valores imposibles, campos vacíos. Consulta los originales si hay dudas.

PASO 4: CLASIFICACIÓN

Agrupar los datos en categorías. Por ejemplo, agrupa por distrito, por tipo de servicio, por grupo de edad, etc.

PASO 5: RESGUARDO

Guarda copias de seguridad de todos los archivos. Nunca trabajes sobre el archivo original; siempre usa una copia.

3.3.5. FASE V: Documentación de los datos

La documentación de datos constituye una fase esencial dentro del ciclo de gestión de la información, cuyo propósito es registrar, describir y organizar de manera estandarizada toda la información generada o utilizada durante los procesos de recolección, procesamiento, análisis y difusión de datos. Este proceso garantiza la trazabilidad, transparencia y reutilización de la información, en concordancia con las directrices emitidas por la Oficina Nacional de Gobierno Electrónico e Informática (ONGEI), las normas sobre Gestión de la Información y Gobierno de Datos en el sector público y otras que determine la Subgerencia de Cooperación Técnica y Planeamiento.

El proceso estará a cargo de la Unidad de Estadística y Gestión del Conocimiento de la Subgerencia de Cooperación Técnica y Planeamiento, asimismo, en coordinación con la Oficina Regional de Desarrollo Institucional y Tecnologías de Información.

a) Identificación del conjunto de datos

Consiste en reconocer y registrar los conjuntos de datos generados, recolectados o utilizados por las unidades orgánicas del GORE Junín. Cada conjunto debe tener un nombre identificador único, una descripción clara de su contenido y estar asociado a un proceso o función institucional.

b) Descripción mediante metadatos

Se debe elaborar una ficha de metadatos para cada conjunto de datos, siguiendo los estándares definidos por la ONGEI y el CEPLAN.

Los metadatos describen el “quién, qué, cuándo, dónde y cómo” de los datos, e incluyen al menos la siguiente información:

- Nombre del conjunto de datos
- Descripción general
- Fuente de origen
- Unidad responsable
- Cobertura temporal y geográfica
- Método de recolección o fuente primaria/administrativa
- Formato del archivo (CSV, XLSX, PDF, etc.)
- Frecuencia de actualización
- Nivel de confidencialidad o acceso (público, restringido o confidencial)
- Fecha de creación o última actualización

c) Clasificación de confidencialidad de datos

De acuerdo con la Ley N.º 29733, Ley de Protección de Datos Personales, y las directivas de la ONGEI, se debe determinar el nivel de confidencialidad de cada conjunto de datos antes de su publicación o intercambio.

Los niveles sugeridos son:

- Público: datos abiertos, accesibles a todos.
- De uso restringido: información institucional no sensible, pero no pública.
- Confidencial: contiene datos personales o estratégicos que requieren medidas de protección o anonimización.

d) Validación técnica y control de calidad

El equipo técnico responsable debe revisar los datos para verificar su consistencia, integridad y exactitud antes de ser documentados o compartidos.

Esta validación puede incluir:

- Revisión de formatos, unidades de medida y nombres de variables.
- Eliminación de duplicidades o errores.
- Confirmación del cumplimiento de estándares técnicos de interoperabilidad definidos por la ONGEI.

e) Almacenamiento y registro en repositorio institucional

Una vez documentados, los datos deben almacenarse en un repositorio institucional digital, organizado por temáticas, años o unidades responsables. Este repositorio permitirá mantener un control sobre la versión vigente de cada conjunto y facilitará su acceso para uso interno y coordinación con otros niveles de gobierno.

3.3.6. FASE VI: Difundir los resultados

Para establecer canales de comunicación eficientes con nuestros grupos de valor y de interés, el GORE Junín debe poner a su disposición la información producida en la entidad, acompañada de la documentación asociada a las operaciones estadísticas realizadas durante el proceso.

Toda información divulgada debe incluir dentro de sus propósitos: facilitar la toma de decisiones de política pública basada en la evidencia, fomentar la cultura de datos al interior de la entidad y empoderar a la ciudadanía, a través de una comunicación con lenguaje claro, de la gestión del Gobierno Regional de Junín, alineada con el índice de transparencia y la política de datos abiertos.

En esta fase, se busca preparar y publicar un conjunto de datos. A nivel del GORE Junín y en el Portal Nacional de Datos Abiertos.

a) Preparación del conjunto de datos

En esta primera fase, las unidades orgánicas que han generado o procesado información consolidan sus conjuntos de datos validados, junto con su documentación técnica (metadatos).

Se revisa que:

- Los datos son consistentes y actualizados.
- La estructura cumple con los formatos aprobados por la ONGEI (CSV, XLSX, JSON, etc.).
- Cada conjunto está acompañado de su ficha de metadatos, que describe contenido, fuente, cobertura, fecha, responsable y nivel de confidencialidad.

Algunos aspectos a tener en cuenta:

- Se mantendrá una copia de respaldo actualizada de la información o base de datos (backup) y trabaje sobre una copia de la información.
- Realice una anonimización de datos sensibles, de ser el caso.

- Elija el formato para la publicación de los datos (CSV, JSON, YAML, XML, otros). Estos deberán seleccionarse conforme a los lineamientos de la ONGEI para garantizar la interoperabilidad.
- Extraiga la información o exporte a un archivo de formato abierto.

b) Validación por el grupo de trabajo de recolección de datos:

El Grupo de Trabajo de Recolección de Datos del GORE Junín realiza la validación técnica y metodológica de la información antes de su publicación.

Las acciones principales en esta fase son:

- Verificar la calidad de los datos (exactitud, integridad, consistencia y pertinencia).
- Confirmar el nivel de confidencialidad, de acuerdo con la Ley N.º 29733.
- Revisar el cumplimiento de los lineamientos institucionales de gestión de información y seguridad digital.
- Validar que los metadatos estén correctamente estructurados según los estándares de la ONGEI y otros que determina la Subgerencia de Cooperación Técnica y Planeamiento.
- Emitir un acta o informe de validación, autorizando la publicación.

c) Selección de tipo de licencia abierta disponible

Las licencias que permiten el uso/reutilización liberando de responsabilidad a las entidades y personas que publican son:

- Public Domain Dedication and License (PDDL)
- Attribution License (ODC-By)
- Open Database License (ODC-ODbL)

En este caso, la Subgerencia de Cooperación Técnica y planeamiento en coordinación con la Oficina Regional de Desarrollo Institucional y Tecnologías de Información serán responsables de seleccionar el tipo de licencia abierta disponible.

d) Publicación de los conjuntos de datos en una herramienta de soporte o repositorio local.

El Gobierno Regional de Junín, de acuerdo a sus recursos, puede optar por el uso de infraestructura propia implementando alguna plataforma o herramienta disponible o contratando un servicio de terceros (en la nube). La publicación inicial se realizará en el repositorio estadístico (no abierta al público), garantizando la seguridad, trazabilidad y respaldo de la información.

e) Publicación de los conjuntos de datos

Los conjuntos de datos aprobados son publicados en los medios oficiales del GORE Junín, garantizando su accesibilidad y trazabilidad. Las opciones incluyen:

- Repositorio institucional de datos del GORE Junín.
- Portal de Transparencia Estándar.
- Portal Nacional de Datos Abiertos (en coordinación con la ONGEI).
- Sitios web temáticos o micrositios de proyectos específicos.

f) Difusión de los datos abiertos publicados

El especialista de comunicación se encargará de la difusión de los datos. En esta fase, se promueve el uso y aprovechamiento de los datos publicados mediante acciones de comunicación institucional, tales como:

- Comunicados oficiales o notas informativas.
- Talleres o capacitaciones sobre el uso de los datos.
- Coordinación con universidades, organizaciones civiles y sectores productivos para fomentar su reutilización.

g) Mantenimiento de los conjuntos de datos

La publicación de datos debe ir acompañada de un sistema de monitoreo continuo para verificar su disponibilidad, actualidad y pertinencia. Publicación de los Datos Abiertos Gubernamentales de la entidad en su plataforma de Datos Abiertos y en el Portal Nacional de Datos Abiertos Gubernamentales.

CAPÍTULO 4:

INDICADORES Y CÓMO MEDIRLOS

Un indicador es como un semáforo que nos dice si algo va bien o va mal. Es una forma de medir el progreso hacia una meta. Por ejemplo, si queremos saber si estamos reduciendo la desnutrición, necesitamos un indicador que nos diga cuántos niños están desnutridos. En este capítulo te explicamos los tipos de indicadores que usamos en el GORE Junín.

4.1 Indicadores de la Cadena de Resultados

Los planes del GORE Junín se organizan en una cadena de resultados: desde los insumos (lo que usamos) hasta los resultados finales (los cambios en la población). En cada eslabón de esta cadena hay indicadores diferentes:

TIPO DE INDICADOR	¿QUÉ MIDE?	EJEMPLOS
INDICADOR DE INSUMO	Mide los recursos que usamos: dinero, personal, equipos, materiales. Responde a la pregunta: ¿Qué tenemos disponible?	• Presupuesto asignado para proyectos • Número de profesionales contratados • Equipos médicos adquiridos
INDICADOR DE PROCESO	Mide las actividades que realizamos: capacitaciones, atenciones, supervisiones. Responde a: ¿Qué estamos haciendo?	• Número de talleres realizados • Atenciones médicas brindadas • Kilómetros supervisados
INDICADOR DE PRODUCTO	Mide los bienes y servicios que entregamos: escuelas construidas, beneficiarios atendidos. Responde a: ¿Qué producimos?	• Escuelas construidas • Productores capacitados • Carreteras rehabilitadas
INDICADOR DE RESULTADO INICIAL	Mide los primeros cambios en la población a corto plazo. Responde a: ¿Qué cambios iniciales vemos?	• % de niños con suplementos de hierro • % de familias con agua potable • % de vías en buen estado

INDICADOR DE RESULTADO INTERMEDIO	Mide cambios de mediano plazo. Requiere esfuerzo de varias entidades. Responde a: ¿Avanzamos hacia la meta?	• % de estudiantes con nivel satisfactorio • % de población con servicios preventivos • % de agricultores con tecnologías sostenibles
INDICADOR DE RESULTADO FINAL	Mide los cambios más importantes de largo plazo. Responde a: ¿Mejoramos la vida de la población?	• Tasa de desnutrición crónica • Tasa de pobreza • Esperanza de vida • Nivel de competitividad regional

4.2 Indicadores de Gestión

Además de los indicadores de resultados, también usamos indicadores de gestión. Estos nos dicen qué tan bien estamos trabajando como institución. Se dividen en cuatro tipos:

TIPO	¿QUÉ MIDE?	EJEMPLOS
EFICACIA	Si estamos cumpliendo nuestras metas. ¿Logramos lo que nos propusimos?	• % de metas del POI cumplidas • % de carreteras rehabilitadas según plan • % de reducción de anemia
EFICIENCIA	Si estamos usando bien los recursos. ¿Logramos más con menos?	• Costo por km de carretera • Costo por paciente atendido • Inspecciones por inspector
CALIDAD	Si lo que hacemos está bien hecho. ¿Están satisfechos los usuarios?	• Tiempo de espera promedio • % de usuarios satisfechos • Tiempo de respuesta a reclamos
ECONOMÍA	Si administramos bien el dinero. ¿Gastamos según lo planificado?	• % de presupuesto ejecutado • % de recuperación de préstamos • % de ingresos propios

CAPÍTULO 5:

¿QUIÉN HACE QUÉ EN LA RECOLECCIÓN?

La recolección de datos no es trabajo de una sola persona. Es un equipo donde cada uno tiene responsabilidades claras. Cuando todos cumplimos nuestro rol, el proceso funciona bien. Aquí te explicamos quién hace qué:

¿QUIÉN?	¿QUÉ HACE EN LA RECOLECCIÓN DE DATOS?
TÚ (El que recolecta)	Eres la base de todo. Sales a campo, haces las encuestas, tomas las mediciones y llenas los formatos. Tu trabajo es asegurar que cada dato que registras sea correcto desde el inicio.
Planificador de tu Unidad Ejecutora	Es el enlace entre los que recolectan y los que consolidan. Recibe los datos de todas las áreas de su unidad, hace una primera revisión y los organiza para enviarlos a la SGCTP.
Jefe de tu Unidad Ejecutora	Es el responsable institucional. Aprueba que los datos de su unidad sean enviados y firma los documentos de validación. Garantiza que su área cumpla con los plazos y estándares.
Subgerencia de Cooperación Técnica (SGCTP)	Dirige todo el proceso. Aprueba los formatos, capacita al personal, consolida los datos de todas las unidades, valida la calidad y publica los resultados en el Observatorio Regional.
Analista Estadístico (SGCTP)	Es el especialista en datos. Revisa la calidad técnica, usa herramientas informáticas para detectar errores, genera indicadores y mantiene actualizado el Observatorio Regional.
Oficina de Tecnologías (ORDITI)	Brinda soporte técnico. Mantiene los sistemas funcionando, implementa plataformas digitales para la captura de datos y publica en el Portal de Datos Abiertos.
Oficina de Comunicaciones	Se encarga de difundir. Comunica a la ciudadanía sobre los datos publicados, elabora materiales de difusión y promueve el uso de la información institucional.

4.7. Conformación del grupo de trabajo para la recolección de datos

Se conformará el Grupo de Trabajo para la Recolección de Datos, el cual estará encargado de coordinar, supervisar y estandarizar los procesos de obtención, sistematización y validación de la información. Su función principal es la elaboración y seguimiento del Plan Anual de Gestión de Datos, que incluirá el inventario de información, las prioridades y las metas de producción estadística. El Grupo se reunirá de manera ordinaria cada tres meses y de forma extraordinaria cuando sea necesario. Sus acuerdos, registrados en actas, serán de cumplimiento obligatorio. La conformación del grupo será formalizada mediante una Resolución Ejecutiva Regional.

La conformación del grupo será formalizada mediante Resolución Ejecutiva Regional emitida por el Titular del GORE Junín o la autoridad que este delegue. Asimismo, estará conformada por:

- **Subgerencia de Cooperación Técnica y Planeamiento:** Unidad encargada del proceso, responsable de la planificación, coordinación y supervisión.
- **Unidades Organizacionales del GORE Junín:** Áreas generadoras de datos, responsables de producir, validar y remitir la información.
- **Oficina Regional de Desarrollo Institucional y Tecnología de la Información:** Encargada de garantizar la compatibilidad técnica, el almacenamiento y la interoperabilidad.
- **Oficina de Comunicaciones:** Responsable de la difusión y promoción pública de la información validada.

Dentro de sus funciones se detallan:

- a) Aprobar el Calendario Anual de Recolección y su ajuste trimestral.
- b) Validar fichas técnicas, formularios, diccionarios de datos y cambios metodológicos.
- c) Resolver inconsistencias o duplicidades entre fuentes administrativas y encuestas.
- d) Calificar la calidad de los datos (índice de validación $\geq 95\%$) antes de su ingreso al repositorio regional.
- e) Priorizar nuevos temas, indicadores o desagregaciones (por ej., género, pueblos indígenas, cambio climático).
- f) Dar vista previa y autorizar la publicación de conjuntos de datos abiertos, asegurando el cumplimiento de la Ley 29733 y derechos de terceros.
- g) Proponer medidas de mejora continua (automatización, nuevas apps móviles, capacitación) y dar seguimiento a su implementación.

4.7.1. Roles del grupo de trabajo

4.7.1.1. Secretario Técnico

El Subgerente de Cooperación Técnica y Planeamiento preside este rol y tiene las siguientes funciones:

- Planificar: Aprueba el calendario anual, las fichas técnicas y el presupuesto de cada operación estadística.
- Coordinar: Convoca y preside la Mesa Regional de Datos, instancia operativa del Grupo de Trabajo que reúne a los generadores de datos, la ORDITI, la Oficina de Comunicaciones y la Unidad de Estadística. Resuelve conflictos de prioridad o recursos.
- Supervisar: Monitorea el avance, aplica sanciones administrativas por incumplimiento y valida el informe final antes de su publicación.
- Garantizar la alineación: Asegura que cada dato recolectado responda a los indicadores del PEI, POI y PDRC, conforme a los lineamientos del CEPLAN.

4.7.1.2. Unidad de estadística y gestión de conocimiento

A cargo del analista estadístico de la Subgerencia de Cooperación Técnica y Planeamiento, sus funciones son:

- Revisión técnica: Valida la consistencia, calidad y metadatos de cada conjunto de datos.
- Registro: Asigna un identificador único y documenta la información en la ficha técnica estándar.
- Gestión del Observatorio Regional: Actualiza dashboards, mantiene el catálogo de indicadores y publica boletines estadísticos.
- Retroalimentación: Comunica observaciones a las unidades generadoras para su corrección.

4.7.1.3. Especialistas informáticos:

Bajo la dirección del titular de la oficina o su delegado, sus responsabilidades incluyen:

- Fijar y actualizar los formatos abiertos prioritarios (CSV, JSON, GeoJSON, RDF/DCAT-AP-PE) para toda publicación de datos del GORE Junín.
- Elaborar el "Catálogo Regional de Datos Abiertos" en línea con el Portal Nacional de Datos.
- Aplicar el protocolo de anonimización (k -anonimidad ≥ 3) y evaluar el riesgo de reidentificación antes de liberar microdatos.
- Calcular trimestralmente el "Indicador de Datos Abiertos" y publicarlo en el Dashboard de Gobierno Abierto.

4.7.1.4. Especialistas en comunicación

Está presidido por el director regional de la Oficina Regional de Comunicaciones o quien delegue, el mismo que contempla las siguientes funciones:

- Diseñar e implementar estrategias de difusión para asegurar el acceso oportuno y comprensible de la información estadística a la ciudadanía.

4.7.1.5. Unidades generadoras de datos

Está presidido por los planificadores o estadistas de las unidades organizacionales del pliego del GORE Junín, el mismo que contempla las siguientes funciones:

- Generar los datos de su competencia en los formatos y plazos establecidos, el mismo día que ocurre el hecho o en el plazo máximo establecido en el calendario regional.
- Aplicar las fichas técnicas y diccionarios de datos definidos, sin crear campos no autorizados.
- Utilizar estándares de codificación vigentes y formatos abiertos cuando se remitan archivos.
- Verificar la coherencia y completitud de los datos antes de su envío.
- Mantener actualizados sus conjuntos de datos y reportar cualquier alteración.
- Anonimizar microdatos sensibles y gestionar los consentimientos informados.
- Participar activamente en las reuniones y talleres convocados por la Secretaría Técnica.
- Reutilizar datos ya publicados para evitar duplicar esfuerzos.
- Sugerir nuevos conjuntos de datos de alto valor para la ciudadanía y priorizarlos en el Plan Anual de Recolección.

CAPÍTULO 6:

DATOS ABIERTOS Y PUBLICACIÓN

7.1 ¿Qué son los Datos Abiertos?

Los datos abiertos son aquellos que el gobierno pone a disposición de todos los ciudadanos de forma gratuita, sin restricciones de uso. Es como una biblioteca pública de información: cualquiera puede entrar, consultar y usar la información para lo que necesite. Por ejemplo, un investigador puede usar los datos para hacer un estudio, un empresario para tomar decisiones de inversión, o un ciudadano para conocer cómo se está gestionando la región.

7.2 Características de los Datos Abiertos

Para que los datos sean considerados 'abiertos', deben cumplir estas características:

CARACTERÍSTICA	¿QUÉ SIGNIFICA?
Públicos	Cualquier persona puede acceder a ellos, no necesitan permiso especial ni registro.
Accesibles	Están disponibles en internet, en formatos que cualquier computadora puede leer.
Gratuitos	No se paga por descargarlos ni por usarlos.
Reutilizables	Se pueden usar para cualquier propósito: estudios, aplicaciones, análisis, etc.
Sin restricciones	No hay límites sobre cómo se pueden usar o compartir.
En formatos abiertos	Se publican en CSV, JSON, GeoJSON, etc., no en formatos propietarios como Excel cerrado.

7.3 El Observatorio Regional de Junín

El Observatorio Regional de Junín es la plataforma donde se publican todos los datos e indicadores de la región. Es como una vitrina donde los ciudadanos pueden ver cómo está nuestra región en diferentes aspectos: educación, salud, economía, medio ambiente, entre otros. El Observatorio tiene cuatro funciones principales:

- Disponibilidad de información: Facilita el acceso público a datos actualizados y confiables.
- Seguimiento de planes: Permite monitorear los avances del PEI, POI, PDRC y otros instrumentos.

- Centro de referencia: Articula fuentes de información institucionales, académicas y territoriales.
- Producción de análisis: Genera informes, boletines y diagnósticos para la toma de decisiones.

7.4 Proceso de Publicación de Datos

Una vez que los datos están validados, siguen este proceso para ser publicados:

PROCESO DE PUBLICACIÓN DE DATOS

PASO 1: PREPARACIÓN

Se revisa que los datos estén completos, actualizados y en el formato correcto (CSV, JSON). Se prepara la ficha de metadatos.

PASO 2: VALIDACIÓN FINAL

El Grupo de Trabajo de Recolección revisa la calidad, confidencialidad y cumplimiento de lineamientos. Emite acta de aprobación.

PASO 3: ANONIMIZACIÓN

Se eliminan o enmascaran los datos personales para proteger la identidad de las personas, cumpliendo la Ley 29733.

PASO 4: PUBLICACIÓN

Los datos se suben al Portal Nacional de Datos Abiertos (datos.gob.pe) y al Observatorio Regional de Junín.

PASO 5: DIFUSIÓN

La Oficina de Comunicaciones informa a la ciudadanía sobre la nueva publicación mediante comunicados, redes sociales y boletines.

CAPÍTULO 9:

DISPOSICIONES FINALES

9.1 Disposiciones Complementarias

1. **Primera disposición complementaria:** Todas las unidades organizacionales deben coordinar con la Subgerencia de Cooperación Técnica y Planeamiento para la producción de bases de datos, a fin de estandarizar instrumentos y metodologías. Los formatos operativos serán aprobados por Resolución de la Gerencia Regional de Planeamiento.
2. **Segunda disposición complementaria:** Las unidades productoras son responsables de la veracidad, completitud y oportunidad de la información. La validación previa será realizada por el responsable de planeamiento de cada unidad.
3. **Tercera disposición complementaria.** La Oficina Regional de Desarrollo Institucional y Tecnologías de la Información brindará el soporte técnico para el registro, almacenamiento y transmisión segura de datos, promoviendo la integración con el Observatorio Regional.
4. **Cuarta disposición complementaria:** El tratamiento de datos personales se efectuará en estricto cumplimiento de la Ley N.º 29733. La Subgerencia de Cooperación Técnica y la Oficina Regional de Desarrollo Institucional garantizarán las medidas de seguridad y confidencialidad.
5. **Quinta disposición complementaria.** La información debe ser centralizada en la Subgerencia de Cooperación Técnica y Planeamiento para su publicación en el portal de datos abiertos y el Observatorio Regional.
6. **Sexta disposición complementaria.** La Gerencia Regional de Planeamiento, Presupuesto y Acondicionamiento Territorial y la Subgerencia de Cooperación Técnica y Planeamiento supervisarán la implementación de estos lineamientos.

9.2 Disposiciones Transitorias

7. **Primera disposición transitoria:** Durante el primer año, las unidades podrán continuar usando sus formatos actuales, pero deberán iniciar el proceso de adecuación progresiva a los nuevos estándares.
8. **Segunda disposición transitoria:** En un plazo máximo de doce (12) meses, se deberá completar el inventario preliminar del 100% de los datos del Gobierno Regional, que servirá de base para el Repositorio Estadístico.
9. **Tercera disposición transitoria:** Durante el primer año, la Subgerencia de Cooperación Técnica y Planeamiento establecerá los protocolos y formatos unificados que serán de uso obligatorio al finalizar el periodo de transición.
10. **Cuarta disposición transitoria:** En los primeros doce (12) meses, la Subgerencia de Cooperación Técnica y Planeamiento organizará talleres de

capacitación sobre la metodología de levantamiento de datos, uso de instrumentos y registro en sistemas.

11. **Quinta disposición transitoria:** Se priorizará la aplicación de los lineamientos en un conjunto de indicadores críticos del PEI y POI como fase piloto para evaluar y ajustar los procedimientos.
12. **Sexta disposición transitoria:** La información recolectada con metodologías anteriores podrá ser incorporada como fuente referencial, previa evaluación de su viabilidad técnica por la Subgerencia de Cooperación Técnica.
13. **Séptima disposición transitoria:** La Subgerencia de Cooperación Técnica y Planeamiento gestionará un convenio de cooperación con la ODEI-Junín para fortalecer la validación técnica de indicadores y desarrollar capacitaciones.

ANEXO

- Anexo N° 01 : Formato de Mapeo de Usuarios del Territorio (Contenido mínimo para su diseño)
- Anexo N° 02 : Formato de Identificación de necesidades información pública
- Anexo N° 03 : Formato de análisis de información pública
- Anexo N° 04 : Formato de inventario de información pública – Catálogos de Datos Abiertos
- Anexo N° 05 : Formato de priorización de información pública
- Anexo N° 06 : Formulario para recolección de datos
- Anexo N° 07 : Guía para determinar la muestra
- Anexo N° 08 : Matriz de registro de incidencias durante la recolección de datos
- Anexo N° 09 : Matriz de trazabilidad de datos (Creación y modificación)
- Anexo N° 10 : Diseño del Observatorio Regional de Junín
- Anexo N° 11 : Flujograma: Gestión integral de datos
- Anexo N° 12 : Flujograma: Recolección de datos de las Unidad Productora
- Anexo N° 13 : Flujograma: Recolección de datos de las Unidades Ejecutoras
- Anexo N° 14 : Flujograma: Recolección de datos basado en encuestas por las unidades ejecutoras
- Anexo N° 15 : Flujograma: Recepción de ficha datos
- Anexo N° 16 : Flujograma: Validación y difusión de datos

Anexo 1: Formato de Mapeo de Usuarios del Territorio (Contenido mínimo para su diseño)

1. Identificación de usuarios externos

CAMPO	DESCRIPCIÓN	EJEMPLO
Nombre o razón social	Nombre de la organización, institución o persona	Cooperativa Agraria "Valle del Mantaro"
Tipo de actor	Marcar una o más opciones	☐ Gobierno local ☐ Organización civil ☐ Empresa privada ☐ Universidad ☐ Comunidad campesina /nativa ☐ ONGs ☐ OTRO: __
Ubicación territorial	Distrito, provincia	Sicaya, Huancayo
Representante de contacto	Nombre y cargo	Juan Pérez – Presidente
Correo / teléfono	Datos de contacto	jperez@gmail.com / 987654321

2. Interés en la producción de datos

CAMPO	DESCRIPCIÓN	EJEMPLO
¿Qué tipo de datos desea que se generen?	Describir brevemente	Rendimiento agrícola por parcela, uso de agua, clima local
¿Con qué fin?	Marcar una o más opciones	☐ Mejora de gestión ☐ Investigación ☐ Incidencia política ☐ Monitoreo comunitario ☐ Comercialización ☐ OTRO: __
¿Está dispuesto a compartir la información que genera su organización?	Sí / No / Bajo condiciones	Sí, si hay reciprocidad o reconocimiento
¿Qué temas le interesan recibir datos de otros?	Elegir temas	☐ Clima ☐ Salud ☐ Educación

CAMPO	DESCRIPCIÓN	EJEMPLO
		☐ Recursos hídricos ☐ Producción ☐ OTRO: ____

3. Capacidad de producción de datos

CAMPO	DESCRIPCIÓN	EJEMPLO
¿Cómo cree recomendable su producción?	Marcar una o más opciones	☐ Encuestas ☐ Entrevistas ☐ Observaciones ☐ manuales ☐ Registros ☐ OTRO: ____
¿Con qué frecuencia se recolecta la información?	Diario / Semanal / Mensual / Eventual	Mensual
¿Qué formato usaría para la recolección de datos?	Excel / PDF / Papel / GPS / OTRO	Excel y GPS

Anexo 2: Formato de identificación de necesidades información pública

Unidad Ejecutora:
Área que reporta:
Responsable:
Contacto:
Fecha de reporte:
Periodo de referencia:

Item	Nombre	Descripción	Responsable	Medio	Formato	Frecuencia	Categoría

CATEGORÍA	SUB CATEGORÍA
Sociodemográfica	▪ Estructura y composición de la población en el territorio ▪ Condiciones en la salud ▪ Condiciones en la educación ▪ Acceso y calidad de los servicios básicos ▪ Estado de pobreza ▪ Seguridad ciudadana ▪ Igualdad de género ▪ Características culturales
Ambiental	▪ Calidad ambiental ▪ Conservación de la biodiversidad biológica ▪ Aprovechamientos sostenibles de los recursos naturales
Económico	▪ Valor bruto de la producción según actividades económicas ▪ Población Económicamente Activa ▪ Corredores económicos ▪ Competitividad económica ▪ Sistema financiero ▪ Tecnología e innovación
Conectividad	▪ Estado de los modos de transporte ▪ Estado y cobertura de la red vial ▪ Servicios y corredores logísticos ▪ Conectividad digital
	▪ Principales peligros y zonas vulnerables ▪ Capacidades institucionales (estimación, prevención, reducción del riesgo de desastres,

	preparación, respuesta, rehabilitación y reconstrucción)
Gobernanza	▪ Presencia del Estado ▪ Modernización pública ▪ Participación ciudadana ▪ Transparencia pública y rendición de cuentas ▪ Gestión fronteriza
Otros	

Anexo 3:
Formato de análisis de información pública

Item	Nombre	Descripción	Responsable	Publicación (Si/No)	Restricción (Si/No)

Anexo 4:
Formato de inventario de información pública – Catálogos de Datos Abiertos

Ítem	Nombre	Descripción	Responsable	Medio	Formato	Frecuencia	Categoría	Publicable (S/N)	Impacto			Dificultad			Total Impacto	Total Dificultad	Plazo
									i1	i2	i3	D1	D2	D3			

- Se determina el Impacto de nivel: i1, i2 e i3, el mismo que se colocará los valores de 1 en caso de que sea SI, de ser NO se coloca el valor de 0 según sea el caso. (i, representa el impacto, D=Dificultad)
- Por su parte, en la dificultad es de nivel: D1, D2 y D3, el mismo que se colocará los valores de 1 en caso de que sea SI, de ser NO se coloca el valor de 0 según sea el caso.
- Para este proceso puede tomar los criterios de evaluación de cada conjunto de datos del paso 1 del Anexo 5.
- Sobre el plazo, se determinan los valores de C=Corto, M=Mediano y L=Largo plazo.

Anexo 5: Formato de priorización de información pública

Item	Nombre	Impacto			Dificultad			Total impacto	Total dificultad	Plazo		
		i1	i2	i3	d1	d2	d3			Corto	Mediano	Largo

Paso 1: Criterios de evaluación a cada conjunto de datos:

DE IMPACTO (i)	DE DIFICULTAD (d)
i1: la información genera valor para la economía: ▪ Crecimiento y competitividad (negocios, empleo, etc.) ▪ Innovación	d1: dificultad para extraer la información de la fuente (archivo, sistema)
i2: la información permite crear valor social ▪ Empoderamiento ciudadano ▪ Control y auditoría social ▪ Participación y compromiso público	d2: dificultad técnica u operativa para poder publicar la información en formato abierto
i3: la información permite crear valor para la gobernanza pública: ▪ Rendición de cuentas ▪ Transparencia ▪ Eficiencia ▪ Anti-corrupción	d3: No se dan las condiciones en la información para cumplir con los 8 principios de los datos abiertos (completos, primarios, oportunos, accesibles, procesables por máquina, no discriminatorios, no propietarios y libres de licencias restrictivas)

Paso 2: Asignación de valor por cada criterio de evaluación de conjunto de datos.

IMPACTO			DIFICULTAD	
Condición presentada	Puntaje		Condición presentada	Puntaje
Alto impacto	3		Mucha dificultad	3
Mediano impacto	2		Mediana dificultad	2
Bajo impacto	1		Poca dificultad	1

Paso 3: Totalizar los valores asignados y grafique el total para determinar el momento de la publicación del conjunto de datos



Anexo 6:
Formulario de Necesidad de Información (Contenido mínimo)

FORMULARIO TIPO PARA RECOLECCIÓN DE DATOS (Usuarios Externos)

N°	Interrogante	Características
1	Datos generales	Tipo de usuario Nombres y apellidos Celular Correo Institución a la que presenta
2	Necesidad de requerimiento de información	Planificación Evaluación Investigación Transparencia
3	Temas de interés	Economía Salud Producción Social Etc
4	¿Qué indicadores o variables?	
5	Desagregación que requiere:	Provincia Distrito Sexo Edad Etnia, etc.
6	¿Con qué frecuencia?	
7	Formato de información:	Excel Shape.file, etc
8	Unidad Orgánica que posiblemente produzca la información	DIRESA DREJ DIREPRO DIRCETUR DRTPE DRVCS Etc.
9	Describa la necesidad:	Descripción breve de la necesidad de requerimiento de información estadística.

FORMULARIO TIPO PARA RECOLECCIÓN DE DATOS (Usuarios Internos)

Será solicitada mediante documento a la Subgerencia de Cooperación Técnica y Planeamiento, el mismo que deberá especificar la necesidad de información y el objetivo de su producción.

Anexo 7:

Guía para determinar la muestra

1. DEFINICIONES

- 1.1. **Universo o población:** Es el conjunto total de elementos, personas o unidades de análisis que poseen una característica común y sobre los cuales se desea obtener información o inferir resultados.
- 1.2. **Muestra:** Es un subconjunto representativo del universo o población, seleccionado mediante métodos estadísticos, con el propósito de obtener información válida y confiable sin necesidad de estudiar a la totalidad de la población.
- 1.3. **Margen de error:** Es el grado de precisión o tolerancia estadística que se acepta entre los resultados obtenidos en la muestra y los valores reales de la población. Expresa el rango máximo en que pueden diferir los resultados y suele expresarse en porcentaje.
- 1.4. **Probabilidad:** Es la probabilidad de que los resultados de la muestra reflejen correctamente las características del universo, dentro del margen de error establecido.

2. SELECCIÓN Y TAMAÑO DE LA MUESTRA

Una muestra es una selección de los elegidos que representan a la población total. El tamaño de la muestra es una porción significativa de la población que cumple con las características del estudio, reduciendo los costos y el tiempo.

El tamaño de la muestra de una encuesta es muy importante para poder realizar el estudio/para el seguimiento y evaluación de las políticas y planes de manera correcta, por lo que hay que tener en cuenta los objetivos y las circunstancias en que se desarrolle la investigación.

1.1. Tamaño de la muestra para una población finita

$$= \frac{N * Z_{\alpha}^2 * p * q}{e^2 * (N - 1) + Z_{\alpha}^2 * p * q}$$

Dónde:

- n= Tamaño de la muestra buscado
- N = Tamaño de la población o universo.
- Z=Parámetro estadístico que depende del nivel de confianza (nc)
- e = Error de estimación máximo aceptado.
- P= Probabilidad de que ocurra el evento estudiado (éxito)
- Q=(1-p)= Probabilidad de que no ocurra el evento estudiado.

1.2. Tamaño de la muestra para una población infinita

Cuando la población total es muy grande ($N > 100,000$) o se desconoce su tamaño exacto, el tamaño muestral puede estimarse mediante la siguiente fórmula simplificada:

$$n = \frac{Z^2 * p * q}{e^2}$$

Dónde:

- n = Tamaño de la muestra buscado
- Z = valor asociado al nivel de confianza (por ejemplo, 1.96 para 95%).
- e = margen de error permitido (en forma decimal; 5% = 0.05).
- P = Probabilidad de que ocurra el evento estudiado (éxito)
- $Q = (1-p)$ = Probabilidad de que no ocurra el evento estudiado.

1.3. Nivel de confianza

El nivel de confianza determina la probabilidad de que los resultados muestrales se encuentren dentro del margen de error respecto al valor real de la población. Los valores más utilizados son los siguientes:

- 90% - Puntuación $Z = 1.645$
- 95% - Puntuación $Z = 1.96$
- 99% - Puntuación $Z = 2.576$

1.4. Margen de error

El margen de error es el límite máximo de diferencia admisible entre el valor estimado a partir de una muestra y el valor real existente en toda la población. Indica cuánta precisión o exactitud deseas que tengan tus resultados.

Generalmente las encuestas se basan en información obtenida de una muestra de la población; es lógico que pueda ocurrir un error de muestreo.

Si el 90% de los encuestados respondió que “Sí”, mientras que el 10% no contestó la encuesta, es un margen de error, podríamos decir “bueno”, a tener un 50-50 o 45-55.

Por lo regular, el margen de error puede ser controlado eligiendo una muestra aleatoria y aumentando el tamaño de la muestra, lamentablemente, el presupuesto puede llegar a ser un limitante.

Menor margen de error requiere un tamaño de muestra más grande. El incrementar el tamaño de la muestra aumenta el nivel de confianza. ¿Qué nivel de confianza se necesita? Las opciones típicas son 90%, 95% o 99%.

3. TIPOS DE MUESTREO

2.1. Muestreo probabilístico

El muestreo probabilístico consiste en seleccionar muestras de una población mediante procedimientos aleatorios, de modo que todos los

elementos tengan la misma probabilidad de ser elegidos. Este tipo de muestreo proporciona resultados más representativos y confiables, ya que reduce el sesgo en la selección.

2.1.1. Muestreo aleatorio simple. Es la técnica más básica del muestreo probabilístico. Todos los miembros de la población tienen la misma probabilidad de ser seleccionados, y la elección se realiza de forma aleatoria, ya sea manualmente o mediante programas informáticos. *Se recomienda usar para poblaciones pequeñas o medianas donde se dispone de un marco muestral completo.*

2.1.2. Muestreo estratificado. Consiste en dividir la población en subgrupos homogéneos llamados estratos, según características relevantes (por ejemplo: edad, sexo, nivel educativo, ámbito geográfico). Luego se selecciona aleatoriamente una muestra de cada estrato de manera proporcional a su tamaño. Se recomienda usar para poblaciones grandes y heterogéneas, en las que se busque garantizar representatividad por grupos.

- Muestreo no aleatorio por accidente. La unidad ejecutora incluye los elementos que le son más convenientes para la muestra.
- Muestreo no aleatorio intencional o de juicio. La idea básica que involucra este tipo de muestra es que la lógica y el sentido común pueden usarse para seleccionar la muestra que sea representativa de una población. Ej. Selección de expertos por el método de experto.
- Muestreo por cuotas: Este se obtiene al especificar las características deseadas de los sujetos que se espera recoger la información y se le deja libertad al investigador para que le aplique los instrumentos necesarios a las personas con esas características. Ej. Se desea hacer un estudio de una población estudiantil de los estudiantes que han repetido el 6. grado y tiene determinada edad o situación en el hogar.
- Determinación de la cantidad de estratos de la población del territorio. $n_0 = \left(\frac{Z}{\epsilon}\right)^2 * p * q$

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Dónde:

- n_0 =Cantidad teórica de elementos de la muestra
- n = Cantidad real de elementos de la muestra a partir de la población asumida o de los estratos asumidos en la población.
- N =Número total de elementos que conforman la población, o número de estratos totales de la población.

- Z = Valor estandarizado en función del grado de confiabilidad de la muestra calculada. Por ejemplo, si se considera trabajar con un 95 % de confiabilidad de la muestra seleccionada, entonces el valor estandarizado a asumir es igual a 1.96. (Para dos colas).

2.1.3. Muestreo por conglomerados. La población se agrupa en conglomerados o grupos naturales (por ejemplo: instituciones, distritos, provincias). Se seleccionan aleatoriamente algunos conglomerados e incluyen todos los elementos que pertenecen a los conglomerados elegidos. Se recomienda usar para poblaciones amplias y dispersas geográficamente, donde no es viable listar a todos los individuos.

2.1.4. Muestreo sistemático. Se seleccionan individuos a intervalos regulares de una lista ordenada. Primero se elige aleatoriamente un punto de inicio, y luego se selecciona cada k -ésimo elemento, donde k es el intervalo obtenido dividiendo el tamaño de la población entre el tamaño deseado de la muestra. Se recomienda usar cuando se dispone de una lista completa y ordenada de la población.

2.2. Muestreo no probabilístico

El muestreo no probabilístico se basa en criterios subjetivos o de conveniencia del investigador, sin emplear métodos aleatorios. Aunque es más rápido y económico, no garantiza la representatividad de toda la población.

2.2.1. Muestreo por conveniencia. El muestreo por conveniencia es una técnica de muestreo no probabilístico que implica la selección de individuos fácilmente disponibles y accesibles. El muestreo por conveniencia es fácil de realizar, pero puede no ser representativo de la población.

2.2.2. Muestreo por cuotas. El muestreo por cuotas es una técnica de muestreo no probabilístico que consiste en seleccionar individuos según cuotas predeterminadas para características demográficas importantes como la edad, el sexo o los ingresos. El muestreo por cuotas es fácil de realizar, pero puede no ser representativo de la población.

2.2.3. Muestreo de bola de nieve. El muestreo de bola de nieve es una técnica de muestreo no probabilístico que consiste en seleccionar individuos basándose en las referencias de otros participantes. Es útil para poblaciones de difícil acceso, pero puede estar sesgado hacia individuos con buenos contactos.

2.2.4. Muestreo intencional. El muestreo intencional es una técnica de muestreo no probabilístico que implica la selección de individuos con

base en criterios específicos, como la experiencia o los conocimientos especializados. El muestreo intencional facilita la investigación que exige un conocimiento específico y carece de representatividad poblacional, mediante la selección estratégica de participantes con base en su experiencia.

4. PASOS PARA SELECCIONAR EL PLAN DE MUESTREO

Un plan de muestreo bien diseñado garantiza la representatividad de las muestras, lo que mejora la validez y la generalización de los hallazgos del estudio para determinar la unidad ejecutora. Estos son los pasos para desarrollar un plan de muestreo:

- 3.1. Definición de los objetivos de la investigación:** El primer paso consiste en definir con claridad los objetivos del estudio o el tipo de información que se requiere recopilar. Una adecuada definición de los objetivos permite a la unidad ejecutora identificar la población de interés, delimitar su alcance y determinar el tipo de información necesaria.
- 3.2. Selección del tamaño de la muestra:** Una vez establecidos los objetivos, se debe calcular el tamaño de la muestra que garantice la representatividad y precisión deseadas.
- 3.3. Elección de la técnica de muestreo:** El siguiente paso consiste en elegir la técnica de muestreo más adecuada, considerando los objetivos del estudio, las características de la población y los recursos disponibles.
- 3.4. Identificación del marco de muestreo:** Una vez definida la técnica, se debe establecer el marco de muestreo, entendido como la lista, registro o base de datos que contiene todos los elementos o unidades que integran la población de estudio.
- 3.5. Cálculo del error de muestreo:** Finalmente, se debe calcular el error de muestreo, entendido como el grado de incertidumbre asociado a las estimaciones obtenidas a partir de la muestra.

Anexo 8:
Matriz de registro de incidencias durante la recolección de datos

N°	Fecha	Lugar o UE	Incidencia detectada	Tipo (operativa / técnica / humana)	Descripción	Impacto en la calidad del dato	Acción inmediata	Responsable
1						Bajo / Medio / Alto		

Anexo 9:
Matriz de trazabilidad de datos (Creación y modificación)

Nº	Fecha	Etapas (creación/modificación/validación)	Responsable	Cambio realizado	Justificación	Evidencia	Versión del dataset
1							

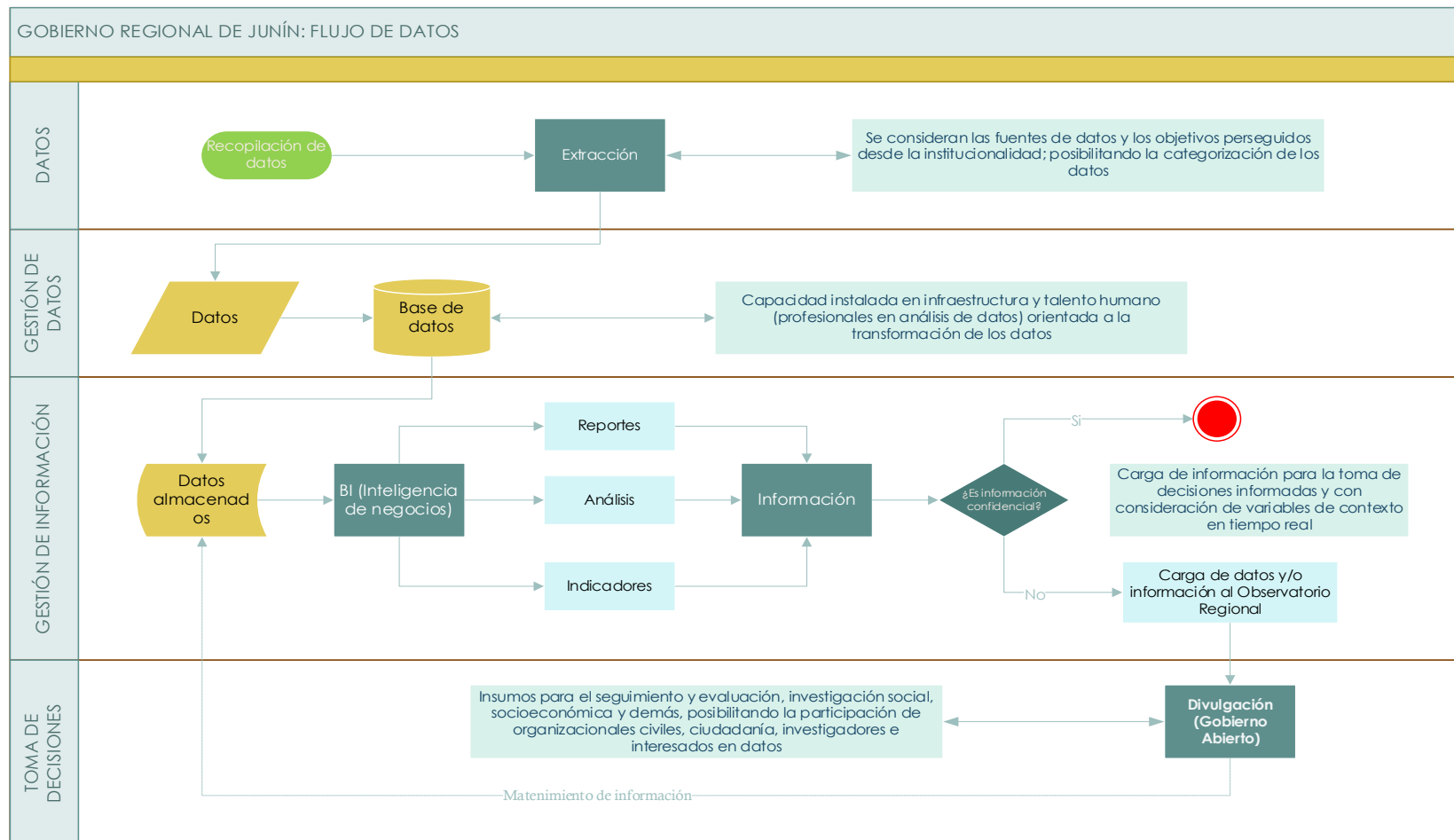
Anexo 10:
Diseño del Observatorio Regional de Junín

IMPLEMENTAR UN OBSERVATORIO REGIONAL

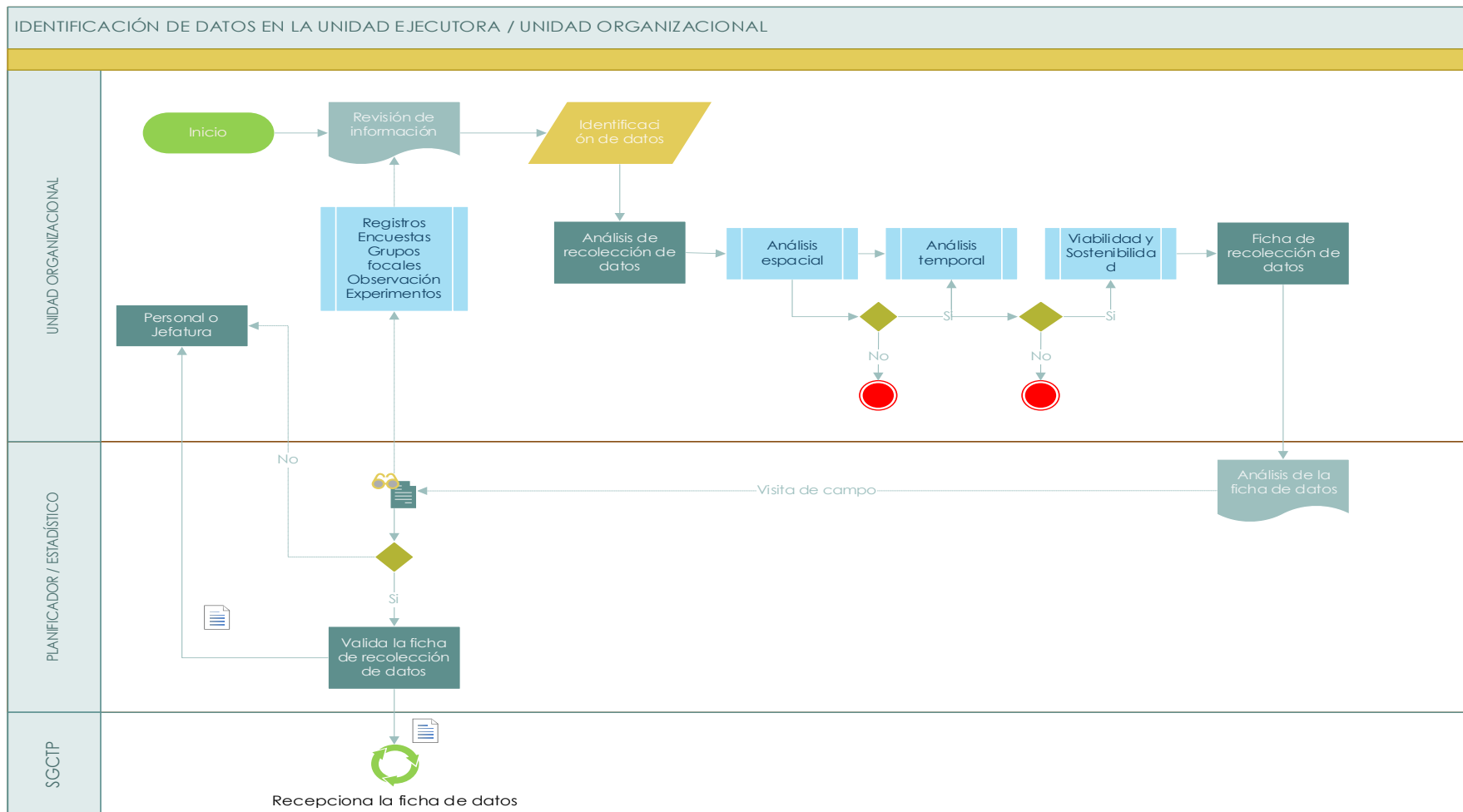
OREJ



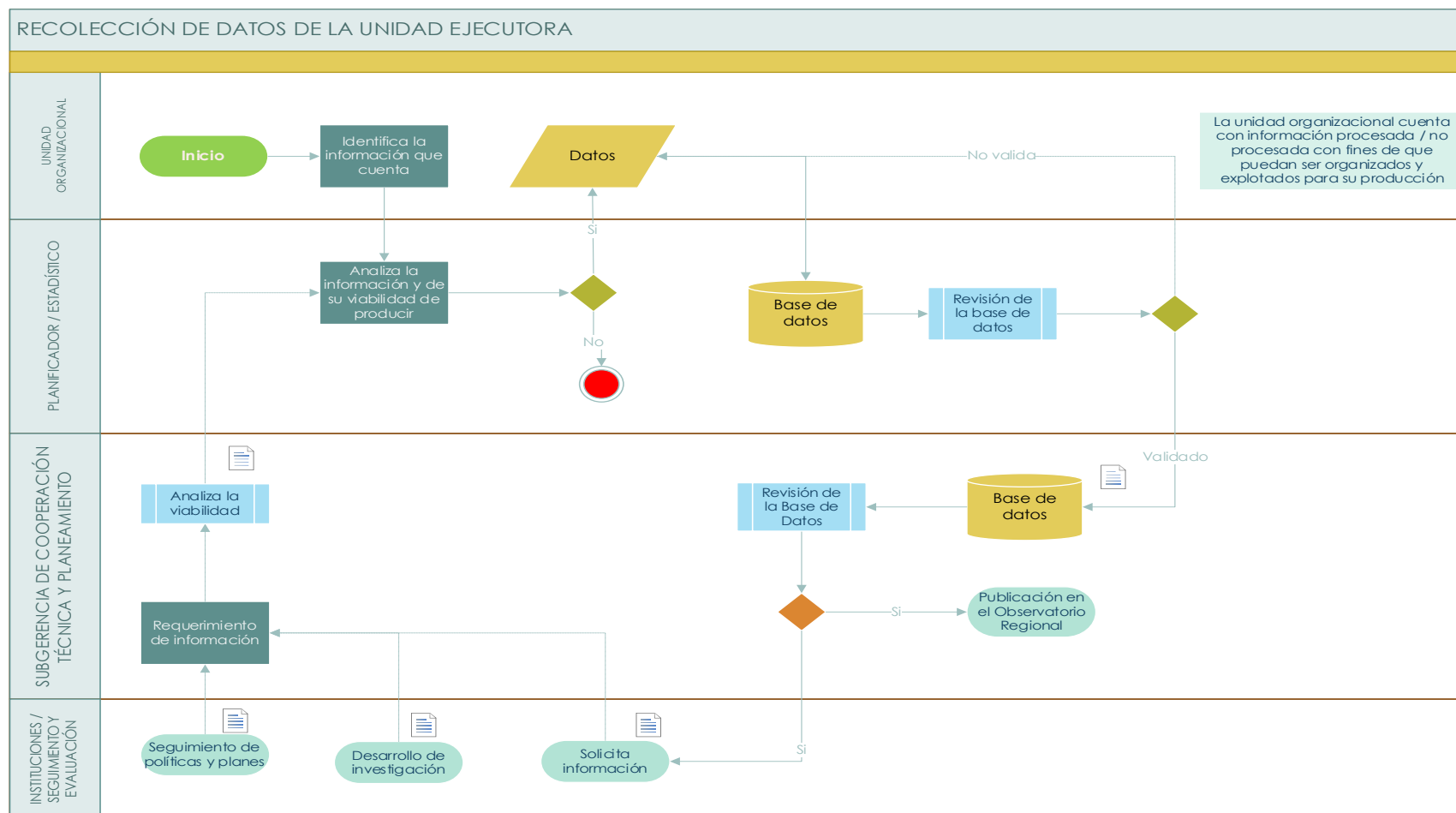
Anexo N° 11 Flujoograma: Gestión integral de Datos



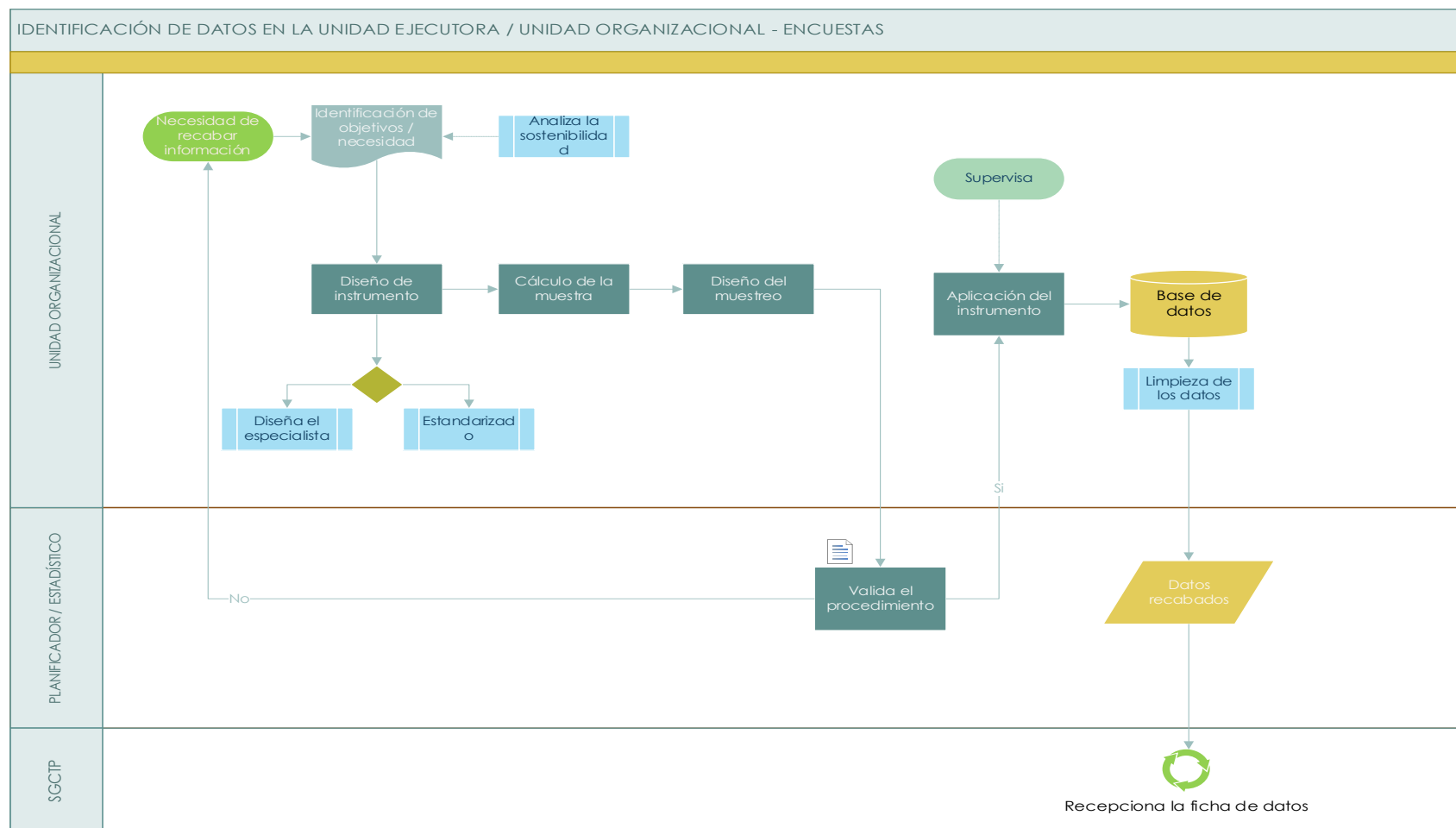
Anexo N°12: Flujograma: Recolección de datos de las Unidad Productora



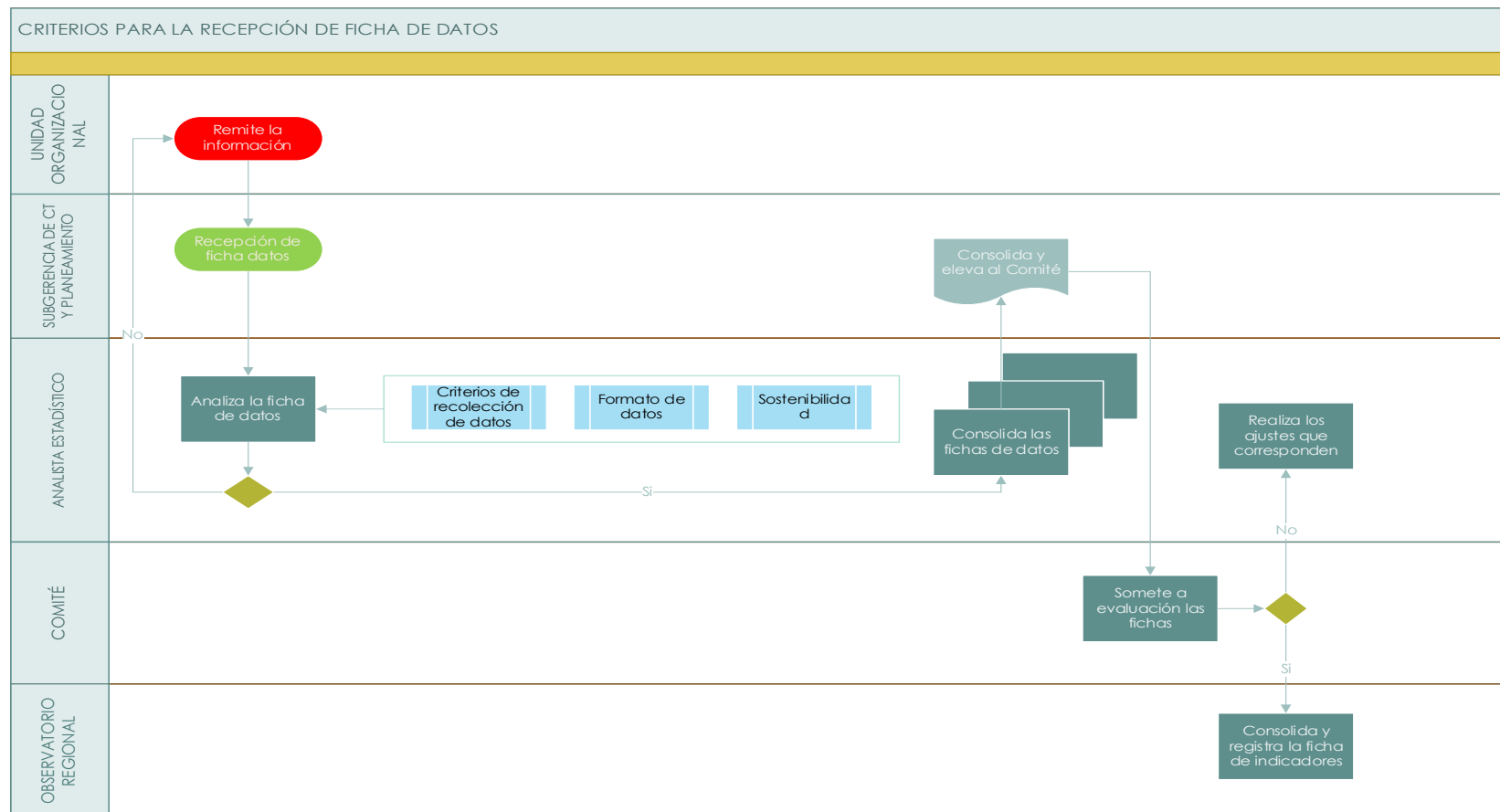
Anexo N° 13: Flujoograma: Recolección de datos de las Unidades Ejecutoras



Anexo N° 14: Flujograma: Recolección de datos basado en encuestas por las unidades ejecutoras



Anexo N° 15: Flujograma: Recepción de ficha datos



Anexo N° 16: Flujograma: Validación y difusión de datos

