

INSTRUCTIVO PARA MINERÍA DE DATOS DE LA INFORMACIÓN ESTADÍSTICA DE INFRAESTRUCTURA Y TRANSACCIONES

GESTIÓN DE ESTUDIOS E INFORMACIÓN DEL SECTOR ELÉCTRICO

MAYO 2021



AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

FIRMAS DE REVISIÓN Y APROBACIÓN: ÁREA DUEÑA DEL PROCESO.

Aprobado por:	Cargo	Firma	Fecha
Santiago Flores	Director de Estudios e Información del Sector Eléctrico		26-05-2021
Revisado por:	Cargo	Firma	Fecha
Andrés Chiles	Especialista de Estudios Eléctricos y Energéticos		26-05-2021
Elaborado por:	Cargo	Firma	Fecha
Christian Junia	Analista de Estudios Eléctricos y Energéticos		26-05-2021
Rodrigo Briones	Profesional		26-05-2021
Diana Cajamarca	Asistente de Estudios Eléctricos y Energéticos		26-05-2021

FIRMAS DE REVISIÓN: OTRAS ÁREAS INCLUIDAS EN EL PROCESO.

Revisado por:	Cargo	Firma	Fecha
N/A	N/A	N/A	N/A


FIRMAS DE REVISIÓN METODOLÓGICA: SISTEMA DE GESTIÓN POR PROCESOS.

Validación metodológica:	Cargo	Firma	Fecha
Paola Cárdenas	Directora de Procesos, Servicios, Calidad, Cambio y Cultura Organizacional		26-05-2021
Asesoría metodológica:	Cargo	Firma	Fecha
Carlos Núñez	Analista de Administración por Procesos 1		26-05-2021

AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

HOJA DE CONTROL DE CAMBIOS

Versión	Pág.	Cambios realizados en el documento	Documento con el que se aprueba	Fecha
01	20	Emisión inicial del documento.	No aplica	15-05-2020
02	18	Se ha actualizado el instructivo de manera integral, en base a la metodología y los formatos desarrollados en la ARC, para el efecto. Esto con la finalidad de alcanzar una automatización completa para la ejecución de la minería de datos de la información estadística de infraestructura y transacciones.	Acta de aprobación N° 0038	26-05-2021

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

1. Objetivo

Establecer las actividades y parámetros técnicos para la ejecución de la minería de datos de la información estadística de infraestructura y transacciones reportada por los Integrantes SISDAT a la Agencia de Regulación y Control de Energía y Recursos Naturales no Renovables - ARC, a través del aplicativo SISDAT.

2. Alcance

El presente instructivo comprende las actividades y parámetros técnicos para la ejecución de la minería de datos de la información estadística de infraestructura y transacciones, desde la captación de la información estadística de infraestructura y transacciones, a través del aplicativo SISDAT, hasta el cierre final de la misma.

3. Lineamientos específicos

- Las actividades relacionadas con el proceso mensual de minería de datos de la información estadística, involucran únicamente a los formularios SISDAT-DEISE utilizados en la GIEG, es decir, los formularios SISDAT de infraestructura y transacciones. Para el efecto, se cuenta con los Especialistas/Analistas Técnicos de la GIEG, mismos que internamente se distribuyen en los siguientes roles, para abarcar los ámbitos relacionados con los formularios SISDAT de infraestructura y transacciones: Especialista/Analista Eléctrico de Infraestructura, Especialista/Analista Eléctrico de Generación y Transmisión, y, Especialista/Analista Eléctrico de Distribución; y, para abarcar aspectos de sistemas de información del SISDAT, el Especialista/Analista de Sistemas.
- Conforme lo establece la *Regulación No. 002/19 Sistema Único de Información Estadística del Sector Eléctrico Ecuatoriano, SISDAT 2.0*, de 10 de diciembre de 2019, el inicio del proceso de minería de datos se efectúa mensualmente a partir del siguiente día laborable de la fecha límite de entrega de la información estadística de infraestructura y transacciones, establecida en el PAO; y, según lo que la Agencia establezca para el efecto, en metodologías, instructivos y procedimientos.
- La entrega y recepción de la información estadística de infraestructura y transacciones se realiza a través del aplicativo SISDAT.


	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

- Con base al numeral 13.1 de la *Regulación No. 002/19 Sistema Único de Información Estadística del Sector Eléctrico Ecuatoriano, SISDAT 2.0*, de 10 de diciembre de 2019, en caso de hallarse inconsistencias (información inexacta o distorsionada) y datos atípicos en la información entregada, detectados en los análisis y procesos inherentes a la minería de datos de la información estadística de infraestructura y transacciones, se notifica este particular al Coordinador/a SISDAT correspondiente, a través de correo electrónico u oficio (despachado desde la CTRCE), y se solicita la ratificación o rectificación de la misma.
- De igual manera, acorde a lo establecido en el mismo numeral de la regulación antes mencionada, en el caso de que se tenga que rectificar la información, los Integrantes SISDAT, en un plazo de cinco (5) días laborables, deben remitir, según criterio de la Agencia: un correo electrónico; o, un informe con los justificativos a las inconsistencias y datos atípicos detectados en la información estadística, mismo que debe ser entregado mediante oficio dirigido al Director/a Ejecutivo, con copia al Coordinador/a Técnico de Regulación y Control Eléctrico.
- En ese sentido, en el caso de que se deba rectificar la información, la Agencia, a través del equipo técnico de la DEISE, brinda asesoría para coordinar y efectuar la rectificación de la información por parte del Integrante SISDAT. Para esto, se pueden efectuar las reuniones que se consideren necesarias para el efecto, de ser el caso, de las cuales se desprenden las actas de reunión correspondientes.
- En el caso de que se ratifique la información, el Integrante SISDAT, en un plazo de cinco (5) días laborables, debe notificar este particular a la ARC mediante correo electrónico u oficio, tal como se señala en la mencionada regulación.

4. Glosario de términos y abreviaturas

4.1. Siglas

- **CTRCE:** Coordinación Técnica de Regulación y Control Eléctrico.
- **DEISE:** Dirección de Estudios e Información del Sector Eléctrico
- **GIEG:** Gestión de Información Estadística y Geográfica.
- **PAO:** Plan Anual de Operación Estadística.
- **SISDAT:** Sistematización de Datos del Sector Eléctrico.

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

4.2. Términos y definiciones

Algoritmos genéticos: Son algoritmos de búsqueda y optimización basados en los mecanismos de selección natural y genética.

Análisis Multivariante: Son técnicas de análisis de datos que estudian la relación simultánea entre el comportamiento de más de dos variables.

Análisis Univariante: Son técnicas de análisis de datos que estudian el comportamiento individual de las variables.

Aplicativo SISDAT: Herramienta informática que apoya a las etapas del proceso de operación estadística del sector eléctrico ecuatoriano. El SISDAT está operando desde el año 2009 y cuenta con información histórica desde el año 1999. SISDAT opera en un ambiente WEB cuya dirección URL es <http://sisdat.controlrecursosyenergia.gob.ec>.


Árbol de decisión: Representación de una función multivariada que fue posible utilizar en la vida práctica a partir del desarrollo de las modernas computadoras. Proveen de una herramienta de clasificación muy potente. Su uso en el manejo de datos la hace ganar en popularidad dadas las posibilidades que brinda y la facilidad con que son comprendidos sus resultados por cualquier usuario.

Base de datos: Es un almacén/repositorio que permite guardar grandes cantidades de información de forma organizada facilitando su ubicación y utilización.

Campo del formulario SISDAT: Corresponde a un parámetro del formulario SISDAT (cabecera o título de una columna).

Captación de la información: Etapa de la operación estadística en la que se ejecutan las actividades de captar (por parte de la ARC) y entregar información estadística generada (por parte de los Integrantes SISDAT), al aplicativo SISDAT.

Cierre final de la información: Etapa de la operación estadística en la que mensualmente se ejecutan instrucciones a nivel de base de datos (SisdatData), con la finalidad de actualizar la información que se consume desde el aplicativo SISDAT, posterior a la ejecución de los análisis y procesos inherentes a la validación y verificación de la consistencia y coherencia de la información estadística entregada por los Integrantes SISDAT.

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

Cierre preliminar de la información: Etapa de la operación estadística en la que mensualmente se ejecutan instrucciones a nivel de base de datos (SisdatData), con la finalidad de actualizar los reportes SISDAT que extraen información desde el aplicativo SISDAT, esto, previo a la ejecución de los análisis y procesos inherentes a la validación y verificación de la consistencia y coherencia de la información estadística entregada por los Integrantes SISDAT.


Coordinador SISDAT: Persona responsable de entregar la información estadística en el SISDAT; y, que tiene a su cargo la coordinación de las tareas necesarias para cumplir con las responsabilidades establecidas en el PAO y en la Regulación correspondiente.

Dato atípico: Es una observación numéricamente distante del resto de los datos, detectado en la etapa de análisis de información de la operación estadística. Es un valor que está fuera del rango normal establecido.

Formulario SISDAT: Documento digital que se utiliza en el aplicativo SISDAT para la recolección de datos de manera estructurada. Puede ser un archivo Excel o un archivo de texto plano en formato .txt.

Formulario SISDAT - DEISE: Son aquellos formularios SISDAT que son gestionados por la DEISE y se utilizan en la GIEG. Los formularios SISDAT-DEISE se dividen en dos grupos: infraestructura y transacciones. Los formularios SISDAT de transacciones se subdividen a su vez en tres subgrupos: transacciones de generación, transmisión y distribución. Los formularios SISDAT de infraestructura son: INF-010 (Caudales), INF-020 (Cobertura Eléctrica), INF-040 (Redes Media Tensión), INF-050 (Luminarias), INF-060 (Redes Secundarias), INF-070 (Acometidas), INF-080 (Medidores) e INF-090 (Personal). Los formularios SISDAT de transacciones de generación son: TRA-010 (Energía Vendida), TRA-180 (Energía Producida) y TRA-190 (Balance de Producción). Los formularios SISDAT de transacciones de transmisión son: TRA-020 (SNT Energía Transportada), TRA-030 (SNT Balance de Energía), TRA-050 (SNT Potencia Línea Transmisión), TRA-090 (SNT DMax Transformadores) y TRA-140 (SNT KV Subestaciones). Los formularios SISDAT de transacciones de distribución son: TRA-040 (Balance Energía), TRA-060 (Pérdidas), TRA-080 (Facturación Clientes No Regulados), TRA-130 (Facturación Clientes Regulados), TRA-160 (Energía Comprada), y TRA-210 (Catastro de Clientes).

Grubbs' Test: También conocida como prueba residual normalizada máxima, es una prueba utilizada para detectar valores atípicos en un conjunto de datos univariado provenientes de una población normalmente distribuida.

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

Inconsistencia: Información estadística inexacta y/o distorsionada, detectada en la etapa de análisis de información de la operación estadística.

Integrante SISDAT: Participantes y/o instituciones del sector eléctrico, CENACE y en general todas las entidades y organismos (públicos o privados) que deben entregar y que consumen información del aplicativo SISDAT.

K-means: Es un algoritmo de clasificación no supervisada (clusterización) que agrupa objetos en k grupos basándose en sus características. El agrupamiento se realiza minimizando la suma de distancias entre cada objeto y el centroide de su grupo o clúster. Se suele usar la distancia cuadrática.


Método bayesiano: Constituye un método elegante y potente, basado en la probabilidad y la estadística que permite construir modelos de probabilidad conjunta manejables que representan las dependencias relevantes entre un conjunto formado por cientos de variables en aplicaciones prácticas. Los modelos resultantes permiten realizar inferencia probabilística de una manera eficiente.

Minería de datos: Es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.

Participantes del sector eléctrico: Personas jurídicas dedicadas a las actividades de generación, autogeneración, transmisión, distribución y comercialización, alumbrado público general, importación y exportación de energía eléctrica, así como también las personas naturales o jurídicas que sean considerados consumidores o usuarios finales.

Plan Anual de Operación Estadística - PAO: Documento en el que se establecen las acciones y requerimientos para el funcionamiento de las etapas de la operación estadística del sector eléctrico; donde se determinan los plazos previstos, módulos de información, formatos, fecha de publicación, responsable que atenderá los requerimientos, mecanismos de captación, entre otros.

Puntuación Z: En estadística, la puntuación Z (o puntuación estándar) de una observación es el número de desviaciones estándar que hay por encima o por debajo de la media de población.

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

Red Neuronal: Conjunto de elementos de procesamiento de la información altamente interconectados, que son capaces de aprender con la información que se les alimenta.

Regresión lineal: En estadística la regresión lineal o ajuste lineal es un modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente Y, las variables independientes Xi y un término aleatorio ϵ .

Reportes SISDAT: Archivos Excel que contienen tablas dinámicas que extraen información desde los formularios SISDAT-DEISE. Se emplean para realizar el proceso manual de análisis de datos de los formularios SISDAT-DEISE que no pueden ser analizados con el Software de Minería de Datos, es decir el INF-010 (Caudales), INF-090 (Personal), TRA-020 (SNT Energía Transportada), TRA-030 (SNT Balance de Energía), TRA-050 (SNT Potencia Línea Transmisión), TRA-060 (Pérdidas), TRA-090 (SNT DMax Transformadores), TRA-130 (Facturación Clientes Regulados), TRA-140 (SNT KV Subestaciones) y TRA-210 (Catastro de Clientes).

Series temporales: Colección de observaciones de una variable recogidas secuencialmente en el tiempo. Estas observaciones se suelen recoger en instantes de tiempo equiespaciados. Si los datos se recogen en instantes temporales de forma continua, se debe o bien digitalizar la serie, es decir, recoger sólo los valores en instantes de tiempo equiespaciados, o bien acumular los valores sobre intervalos de tiempo.

Software de Minería de Datos: Software utilizado para la ejecución de la minería de datos de la información estadística.

5. Descripción de tareas

N°	Responsable	Tarea	Descripción	Documento que genera la tarea
1.	Especialista/Analista Técnico de la GIEG	Receptar información estadística de infraestructura y transacciones	Se recepta la información estadística de infraestructura y transacciones del sector eléctrico, a través del aplicativo SISDAT. Con la información disponible en dicho aplicativo, se procede a ejecutar el cierre preliminar de la	Información estadística de infraestructura y transacciones

<div>AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES</div> <div><div></div><div></div><div></div></div>	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

			misma.	
2.	Especialista/Analista de Sistemas de la GIEG	Ejecutar cierre preliminar de la información estadística de infraestructura y transacciones	Se ejecutan instrucciones a nivel de base de datos (SisdatData), con la finalidad de actualizar los reportes SISDAT. Para este propósito, se utiliza la herramienta Microsoft SQL Server, en la cual se ejecuta el archivo "cierreMensual.sql", ubicado en la dirección: M:\CNRSE\DNEEE\ESTADISTICAS\SQL\cierre.	
3.	Especialista/Analista Técnico de la GIEG	Preparar ejecución de minería de datos	<p>La ejecución de la minería de datos de la información estadística de infraestructura y transacciones comprende dos procesos en paralelo. El primero hace uso del Software de Minería de Datos; y, el segundo, se efectúa mediante un proceso manual de análisis de datos utilizando reportes SISDAT (para aquellos formularios SISDAT-DEISE que no pueden ser analizados con el Software de Minería de Datos).</p> <p>Proceso con Software de Minería de Datos:</p> <p>Con este software se realiza la minería de datos de los formularios SISDAT-DEISE:</p> <p>INF-020 (Cobertura Eléctrica), INF-040 (Redes Media Tensión), INF-050 (Luminarias), INF-060 (Redes Secundarias), INF-070 (Acometidas), INF-080 (Medidores), TRA-010 (Energía Vendida), TRA-040 (Balance Energía), TRA-080 (Facturación Clientes No Regulados), TRA-160 (Energía Comprada), TRA-180 (Energía Producida), y,</p>	

<div>AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES</div> <div></div>	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

			TRA-190 (Balance de Producción). Para iniciar la ejecución de la minería de datos con este software: - Se ingresa al software. Se ingresa Usuario y Clave. - Se selecciona el comando “Ejecutar”, luego se selecciona “Procesos”. Proceso manual - Reportes SISDAT: Se eligen y abren los reportes SISDAT para iniciar el proceso manual de análisis de datos de los formularios SISDAT-DEISE siguientes: INF-010 (Caudales), INF-090 (Personal), TRA-020 (SNT Energía Transportada), TRA-030 (SNT Balance de Energía), TRA-050 (SNT Potencia Línea Transmisión), TRA-060 (Pérdidas), TRA-090 (SNT DMax Transformadores), TRA-130 (Facturación Clientes Regulados), TRA-140 (SNT KV Subestaciones), y, TRA-210 (Catastro de Clientes).	
4.	Especialista/Analista Técnico de la GIEG	Elegir parámetros necesarios para ejecución de minería de datos	Proceso con Software de Minería de Datos: Se elige el tipo de análisis: Univariante o Multivariante. Para cada técnica de análisis se elige el tipo de algoritmo a emplear. Para el caso del análisis univariante los algoritmos a emplear son: GrubbTest, Puntuación Z, Regresión Lineal,	

AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	


		<p>Series Temporales.</p> <p>Redes Neuronales***,</p> <p>Algoritmos genéticos***.</p> <p>Para el caso del análisis multivariante los algoritmos a emplear son:</p> <p>K-means,</p> <p>Árboles de decisión***,</p> <p>Métodos bayesianos***.</p> <p>Se elige el formulario SISDAT-DEISE al que se realizará el análisis:</p> <p>INF-020 (Cobertura Eléctrica),</p> <p>INF-040 (Redes Media Tensión),</p> <p>INF-050 (Luminarias),</p> <p>INF-060 (Redes Secundarias),</p> <p>INF-070 (Acometidas),</p> <p>INF-080 (Medidores),</p> <p>TRA-010 (Energía Vendida),</p> <p>TRA-040 (Balance Energía),</p> <p>TRA-080 (Facturación Clientes No Regulados),</p> <p>TRA-160 (Energía Comprada),</p> <p>TRA-180 (Energía Producida),</p> <p>TRA-190 (Balance de Producción),</p> <p>INF-010 (Caudales)**,</p> <p>INF-090 (Personal)**,</p> <p>TRA-020 (SNT Energía Transportada)**,</p> <p>TRA-030 (SNT Balance de Energía)**,</p> <p>TRA-050 (SNT Potencia Línea Transmisión)**,</p> <p>TRA-060 (Pérdidas)*,</p> <p>TRA-090 (SNT DMax Transformadores)**,</p> <p>TRA-130 (Facturación Clientes Regulados)*,</p> <p>TRA-140 (SNT KV Subestaciones)** y,</p>	
--	--	--	--

<div>AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES</div> <div></div>	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

			TRA-210 (Catastro de Clientes)*. Finalmente se elige el año, el mes y la empresa (Integrante SISDAT). Proceso manual - Reportes SISDAT: Se actualizan las tablas dinámicas de los reportes SISDAT. Se elige el año, el mes, la empresa (integrante SISDAT), y el campo del formulario SISDAT-DEISE a analizar (análisis univariante).	
5.	Especialista/Analista Técnico de la GIEG	Ejecutar minería de datos	Proceso con Software de Minería de Datos: Una vez elegidos los parámetros, se ejecuta el comando “Procesar”. Con la ejecución de este comando, el software realiza la detección automática de inconsistencias y datos atípicos. Cuando se quiere procesar información que ya ha sido procesada previamente, el software presenta el siguiente mensaje: “¡Ya existen datos procesados!. ¿Desea volver a procesarlos?”; y, se da la opción de “Si” o “No”. Si se elige “Si” el software vuelve a procesar la información. Si se elige “No”, el software permite consultar los resultados previamente procesados, mediante la ejecución del comando “Consultar”. Proceso manual - Reportes SISDAT: Se realiza la detección de inconsistencias y datos atípicos, mediante el análisis de la varianza o desviación estándar de los datos de cada campo de los formularios SISDAT-DEISE	

AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

			<p>que se analizan mediante reportes SISDAT. Para este propósito se emplea las herramientas integradas de Excel como tablas dinámicas, macros, power pivot, gráficos de tendencias, entre otras.</p> <p>Con base a los reportes SISDAT de transacciones de generación, se efectúan análisis comparativos de los datos históricos, fijando como límites la producción mensual de energía a potencia nominal y un precio medio de 40 USD¢/kWh.</p>	
6.	Especialista/Analista Técnico de la GIEG	Generar resultados de minería de datos	<p>Proceso con Software de Minería de Datos:</p> <p>Los resultados de la minería de datos se los exporta en un archivo de Excel mediante la ejecución del comando “Exportar”. El archivo (ver Anexo 7.1) consta de dos pestañas: la primera presenta un resumen por empresa (Integrante SISDAT) de las anomalías encontradas en la ejecución de la minería de datos; mientras que en la segunda pestaña se encuentra el detalle de las anomalías. Este software no permite obtener resultados gráficos de la minería de datos como los obtenidos mediante los reportes SISDAT (Ver Anexo 7.2***).</p> <p>Proceso manual - Reportes SISDAT:</p> <p>Los resultados de la detección de inconsistencias y datos atípicos, se los presenta mediante tablas y gráficos realizados manualmente con las herramientas integradas de Excel (ver Anexo 7.2).</p>	Resultados de la minería de datos

	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

			¿Existen inconsistencias y datos atípicos? SI: Continúa en la actividad 7. NO: Continúa en la actividad 10.	
7.	Especialista/Analista Técnico de la GIEG	Reportar inconsistencias y datos atípicos	Se notifica al Coordinador SISDAT las inconsistencias y datos atípicos, detectados en los análisis y procesos inherentes a la minería de datos de la información estadística de infraestructura y transacciones; y, se solicita la ratificación o rectificación de la misma, según lo establecido en la Regulación 002/19.	Oficio / Correo electrónico
8.	Especialista/Analista Técnico de la GIEG	Receptar y analizar respuesta del Integrante SISDAT	Recepta y analiza la respuesta por parte del Integrante SISDAT, en la que ratifica la información o solicita la rectificación de la misma, según lo establecido en la Regulación 002/19. ¿Se debe rectificar la información estadística? SI: Continúa en la actividad 9. NO: Continúa en la actividad 10.	Oficio / Correo electrónico
9.	Especialista/Analista Técnico de la GIEG	Gestionar rectificación de inconsistencias y datos atípicos en la información estadística	En el caso de que se deba rectificar la información estadística, coordina con el Integrante SISDAT, brinda asesoría, se llevan a cabo reuniones de trabajo, de ser el caso, y se efectúa la rectificación de la información en el aplicativo SISDAT.	Información estadística rectificada
10.	Especialista/Analista de Sistemas de la GIEG	Ejecutar cierre final de información estadística de infraestructura y	Una vez que se haya validado y verificado la información disponible en el aplicativo SISDAT, se ejecutan instrucciones a nivel de base de datos, con la finalidad de actualizar los reportes SISDAT que extraen información desde el	Base de datos SISDAT actualizada y depurada

AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES 	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

		transacciones	aplicativo SISDAT.	
			FIN DEL PROCESO	

Nota: Por limitantes técnicas, ciertas sub actividades no se están realizando o se han dejado de realizar, por lo que en el presente documento se han establecido grados de criticidad de ejecución de dichas subactividades con los siguientes criterios:

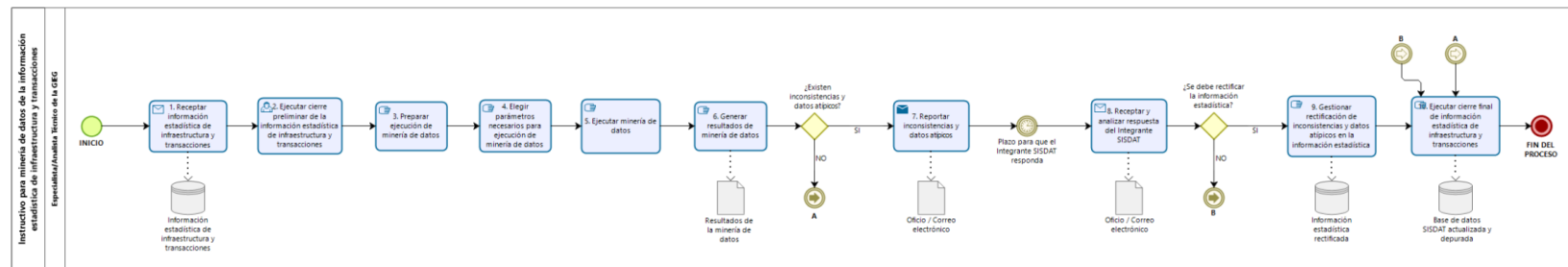
(*) Muy necesario,

(**) Necesario,

(***) No tan necesario.

AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

6. Diagrama de flujo



AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES	Instructivo para minería de datos de la información estadística de infraestructura y transacciones	Código: GTRCE.GEISE.03.IT.01
	Macro proceso: Gestión Técnica de Regulación y Control Eléctrico	Acta N°: 0038
Fecha de Aprobación: 26-05-2021	Proceso: Gestión de Estudios e Información del Sector Eléctrico	Versión: 02
	Subproceso: Gestión de la información estadística del sector eléctrico	

7. Anexos

7.1. Formato de Resultados de la minería de datos obtenidos mediante Software de Minería de Datos


7.1.1. Primera pestaña: Anomalías por Agente

Empresa	Total atípicos	Unidad
Empresa 1		celdas
Empresa 2		celdas
Empresa 3		celdas
Empresa 4		celdas

7.1.2. Segunda pestaña: Detalle de Anomalías

Agente	Año	Mes	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable n
Empresa 1	n	m								
Empresa 2	n	m				(Dato atípico)				
Empresa 3	n	m								
Empresa 4	n	m		(Dato atípico)					(Dato atípico)	

7.2. Formato de Resultados de la minería de datos obtenidos mediante reportes SISDAT

 <div>AGENCIA DE REGULACIÓN Y CONTROL DE ENERGÍA Y RECURSOS NATURALES NO RENOVABLES</div>		<div>Formato de Resultados de la minería de datos obtenidos mediante Reportes SISDAT</div> <div>GTRCE GEISE.03.FO.01</div> <div>Acta de Aprobación N° 0038 Fecha Aprobación: 26-05-2021</div>														Versión: 01	
Empresa Año		<div>CHL Bolivia</div>															
DATOS SISDAT		<div>Validar SISDAT</div>															
Parámetros de Control		Unidad	dic-19	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	ene-21	Tendencia
Longitud Total de Líneas de Media Tensión		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Número Total de Transformadores		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MVA Totales de Transformadores		MVA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Longitud Total de Líneas de Baja Tensión		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cantidad Total de Luminarias		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Potencia Total de Luminarias		kW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cantidad Total de Acometidas		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Longitud Total de Acometidas		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cantidad Total de Medidores		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DATOS GDB		<div>Validar GIS</div>															
Característica		Unidad	dic-19	Ene	Feb	Mar*	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	ene-21	Tendencia
Longitud Total de Líneas de Media Tensión		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Longitud Total de Líneas de Baja Tensión		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cantidad Total de Luminarias		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cantidad Total de Acometidas		#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Longitud Total de Acometidas		km	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Nota: Datos atípicos se resaltan en las tablas para validaciones y rectificaciones