

Descripción de tablas y campos en Oracle del Data Warehouse

Este documento describe la segunda versión del modelo conceptual del Data Warehouse de cada una de las tablas utilizadas en el modelo híbrido para datos de personas y empresas desarrollado para el motor de base de datos Oracle, su propósito y la descripción detallada de cada uno de sus campos. El modelo se construyó bajo un enfoque moderno y modular en Oracle con la creación de un esquema llamado SIRE_USER en el cual va a contener todas las tablas relacionadas con del Data Warehouse pero divididas en capas.

Capa raw (datos crudos)

Contiene los datos tal como se reciben de las fuentes originales, sin transformaciones. Esta capa garantiza:

- Trazabilidad directa con el origen.
- Capacidad de reprocesamiento.
- Auditoría de versiones y control de calidad.

Capa staging (modelo Data Vault 2.0)

Transforma e integra los datos brutos en estructuras hub-satélite-link, siguiendo principios de Data Vault. Permite:

- Manejo eficiente del cambio.
- Conservación del historial completo.
- Separación clara entre claves de negocio y atributos descriptivos.
- Alta escalabilidad y flexibilidad.

3. Capa BI (modelo estrella o dimensional)

Expone los datos integrados y depurados en un modelo dimensional optimizado para analítica, que facilita:

- La construcción de dashboards, reportes y consultas eficientes.
- El análisis histórico de dimensiones y hechos clave.
- La comprensión por parte de usuarios no técnicos.

Nota: El Data Warehouse implementado en Oracle incluye **tablas de catálogos** que desempeñan un papel fundamental en la normalización y estandarización de los datos.

Tabla de contenido

Catálogos:	3
Tabla: catalogo_fuente_registro.....	3
Tabla: catalogo_tipo_documento.....	3
Tabla: catalogo_pais.....	3
Tabla: catalogo_departamento.....	4
Tabla: catalogo_municipio.....	4
Tabla: catalogo_ubicacion_geografica.....	5
Tabla: catalogo_ciiu.....	5
Capa: raw:	6
Tabla: raw_log_etl.....	6
Tabla: raw_bdua.....	6
Tabla: raw_rues.....	7
Capa: staging:	9
Tabla: staging_hub_persona.....	9
Tabla: staging_sat_persona.....	9
Tabla: staging_hub_entidad.....	10
Tabla: staging_sat_entidad.....	10
Tabla: staging_hub_documento.....	10
Tabla: staging_sat_documento.....	11
Tabla: staging_link_persona_documento.....	11
Tabla: staging_link_migracion_persona.....	11
Tabla: staging_link_nacimiento_persona.....	12
Tabla: staging_link_fallecido_persona.....	12
Tabla: staging_link_estado_civil_persona.....	13
Tabla: staging_hub_empresa.....	13
Tabla: staging_sat_empresa.....	14
Tabla: staging_link_representante_legal_empresa.....	15
Tabla: staging_link_empresa_actividad_economica.....	15
Tabla: staging_link_persona_afiliacion.....	16
Capa: bi:	16
Tabla: bi_dim_tiempo.....	16
Tabla: bi_dim_ubicacion_geografica.....	17
Tabla: bi_dim_persona.....	17
Tabla: bi_dim_documento.....	18
Tabla: bi_dim_estado_civil.....	18
Tabla: bi_fact_afiliacion.....	19
Tabla: bi_fact_migracion_persona.....	19
Tabla: bi_fact_fallecimiento_persona.....	20
Tabla: bi_fact_nacimiento_persona.....	20
Tabla: bi_fact_estado_civil_persona.....	20
Tabla: bi_dim_empresa.....	21

Catálogos:

Tabla: **catalogo_fuente_registro**

Describe las fuentes oficiales desde donde se originan los datos utilizados en el sistema, tales como entidades gubernamentales, sistemas externos o plataformas institucionales.

Campos:

- **id**: Identificador único de la fuente de datos.
- **descripcion**: Descripción textual y funcional de la fuente registrada.
- **sistema_origen**: Sistema, plataforma o entidad de donde provienen los datos.
- **tipo_fuente**: Tipo de fuente (ej. archivo plano, API, sistema en línea).
- **fecha_creacion**: Fecha en que se registró el dato en el sistema.
- **usuario_carga**: Usuario responsable de registrar la fuente.
- **estado**: Estado actual de la fuente (**activo**, **inactivo**, etc.).

Tabla: **catalogo_tipo_documento**

Catálogo de tipos de documento de identidad utilizados para personas naturales y jurídicas.

Campos:

- **id**: Identificador único del tipo de documento.
- **abreviacion_documento**: Abreviación oficial del documento (ej. CC, TI).
- **nombre_documento**: Nombre completo del documento (ej. Cédula de Ciudadanía).
- **descripcion**: Descripción extendida del tipo de documento.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **catalogo_pais**

Catálogo de países con sus respectivos códigos internacionales y datos geográficos.

Campos:

- **id**: Identificador único del país.
- **codigo_pais**: Código nacional o DIVIPOLA del país.
- **nombre_es**: Nombre del país en idioma español.
- **nombre_fr**: Nombre del país en idioma francés.
- **alpha2**: Código ISO de dos letras (Alpha-2).
- **alpha3**: Código ISO de tres letras (Alpha-3).

- **latitud**: Coordenada geográfica de latitud.
- **longitud**: Coordenada geográfica de longitud.
- **estatus**: Estado del proceso.
- **load_datetime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **catalogo_departamento**

Catálogo de departamentos administrativos del país con datos geográficos de referencia.

Campos:

- **id**: Identificador único del departamento.
- **codigo_departamento**: Código oficial del departamento (DIVIPOLA).
- **nombre**: Nombre del departamento.
- **latitud**: Coordenada geográfica de latitud del centroide del departamento.
- **longitud**: Coordenada geográfica de longitud del centroide del departamento.
- **load_datetime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **catalogo_municipio**

Catálogo de municipios del país con datos geográficos y su correspondiente relación al departamento.

Campos:

- **id**: Identificador único del municipio.
- **codigo_municipio**: Código oficial del municipio (DIVIPOLA).
- **nombre**: Nombre completo del municipio.
- **latitud**: Coordenada geográfica de latitud del municipio.
- **longitud**: Coordenada geográfica de longitud del municipio.
- **id_departamento**: Identificador del departamento al que pertenece.
- **load_datetime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **catalogo_ubicacion_geografica**

Tabla que consolida direcciones geográficas normalizadas, con detalles de ubicación asociada a municipios, departamentos y zonas.

Campos:

- **id**: Identificador único del registro de ubicación.\
- **llave_direccion**: Código o clave única asociada a una dirección o ubicación.
- **direccion**: Dirección detallada en texto libre.
- **id_municipio**: Identificador del municipio relacionado.
- **id_departamento**: Identificador del departamento relacionado.
- **tipo_zona**: Tipo de zona donde se ubica (**urbana**, **rural**, **especial**).
- **latitud**: Coordenada de latitud de la ubicación.
- **longitud**: Coordenada de longitud de la ubicación.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **catalogo_ciiu**

Catálogo de códigos CIIU utilizados para clasificar las actividades económicas de empresas y personas naturales según la Clasificación Industrial Internacional Uniforme.

Campos:

- **id**: Identificador único del código CIIU.
- **codigo_ciiu**: Código CIIU de 4 caracteres.
- **descripcion**: Descripción textual de la actividad económica correspondiente al código.

Capa: **raw**

Tabla: **raw_log_etl**

Almacena el historial de ejecuciones de procesos ETL que alimentan el Data Warehouse, permitiendo su trazabilidad y auditoría.

Campos:

- **id**: Identificador único del registro de ejecución ETL.
- **job_uuid**: Código único que identifica la ejecución del proceso.
- **start_time**: Fecha y hora de inicio de la ejecución.
- **end_time**: Fecha y hora de finalización del proceso.
- **estado**: Estado final del proceso.
- **registros_cargados**: Número de registros procesados por la carga.
- **tabla_destino**: Tabla objetivo en la cual se insertaron los datos.
- **mensaje_error**: Texto que describe el error, en caso de presentarse.
- **nombre_etl**: Nombre del proceso ETL.
- **descripcion_etl**: Descripción funcional del proceso ejecutado.
- **tiempo_ejecucion_segundos**: Tiempo total de ejecución del proceso, en segundos.
- **nombre_archivo**: Nombre del archivo fuente utilizado en la carga.
- **hash_archivo**: Código hash del archivo de entrada, para control de integridad.

Tabla: **raw_bdua**

Contiene los registros brutos provenientes de la fuente BDUA (Base de Datos Única de Afiliados del Sistema de Seguridad Social en Salud). Esta tabla almacena información sin transformar, tal como es recibida, y constituye el insumo inicial para los procesos de integración, transformación y modelado posterior en el Data Warehouse

Campos:

- **TIPO_IDENTIFICACION**: Tipo de documento de identidad de la persona afiliada (CC, TI, CE, etc.).
- **NUMERO_IDENTIFICACION**: Número del documento de identidad del afiliado.
- **PRIMER_APELLIDO**: Primer apellido del afiliado.
SEGUNDO_APELLIDO: Segundo apellido del afiliado.
- **PRIMER_NOMBRE**: Primer nombre del afiliado.
SEGUNDO_NOMBRE: Segundo nombre o nombres adicionales del afiliado.
- **FECHA_NACIMIENTO**: Fecha de nacimiento del afiliado.
SEXO: Sexo registrado del afiliado (M, F u otras clasificaciones según norma).
- **Cod_departamento**: Código del departamento (DIVIPOLA) correspondiente a la residencia o afiliación.
- **Cod_municipio**: Código del municipio (DIVIPOLA) correspondiente a la residencia o afiliación.
Zona: Clasificación de la zona geográfica de residencia (**urbana**, **rural**, etc.).

- **Cod_entidad**: Código de la Entidad Administradora de Planes de Beneficio (EPS o similar) que gestiona la afiliación.
- **Nombre_entidad**: Nombre de la entidad administradora (EPS) a la que pertenece el afiliado.
- **Fecha_afiliacion**: Fecha en que se registró la afiliación a la EPS o entidad.
- **Estado_afiliacion**: Estado actual de la afiliación (activo, suspendido, retirado, etc.).
- **Tipo_documento_cotizante**: Tipo de documento del cotizante principal (persona que realiza los aportes).
- **numero_documento_cotizante**: Número del documento del cotizante.
- **Tipo_cotizante**: Clasificación del tipo de cotizante (dependiente, independiente, pensionado, etc.).
- **TIPO_AFILIADO**: Tipo de afiliado (cotizante, beneficiario, adicional, etc.).
- **PARENTESCO**: Relación del afiliado con el cotizante (hijo, cónyuge, madre, etc.).
- **discapacidad**: Indicador de si la persona tiene condición de discapacidad registrada.
- **Tipo_poblacion_especial**: Clasificación si pertenece a población especial (desplazado, indígena, ROM, etc.).
- **Etnia**: Grupo étnico reportado por el afiliado.
- **Nombre_resguardo**: Nombre del resguardo indígena, si la persona pertenece a uno.
- **IPS_PRIMARIA**: Institución Prestadora de Servicios de Salud asignada como principal para el afiliado.
- **Tipo_identificacion_aportante**: Tipo de identificación del aportante (empresa u organización que realiza aportes).
- **numero_identificacion_aportante**: Número del documento de identificación del aportante.
- **Razon_social_aportante**: Razón social o nombre de la empresa aportante.
- **Actividad_economica**: Código de la actividad económica según CIIU asociada al aportante.
- **REGIMEN**: Régimen de afiliación del afiliado (**contributivo**, **subsidiado**, **especial**, etc.).
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_estadistico_persona**: Identificador estadístico único asignado a la persona en el proceso de integración de datos.
- **llave_direccion**: Clave de relación con la ubicación geográfica del domicilio de la persona.
- **persona_hash**: hash de la persona.

Tabla: **raw_rues**

Contiene información general de las empresas registradas en el país, incluyendo datos de identificación tributaria, estado legal, matrícula mercantil, y representante legal.

Campos:

- **id_estadistico_empresa**: Identificador estadístico único de la empresa.
- **llave_direccion**: Clave que relaciona la empresa con su ubicación geográfica.

- **codigo_camara**: Código de la Cámara de Comercio que realiza la actualización de datos.
- **camara_comercio**: Nombre de la Cámara de Comercio correspondiente.
- **matricula**: Número de Matrícula Mercantil asignada por la Cámara de Comercio.
- **inscripcion_proponente**: Número de inscripción en el registro de Proponentes.
- **razon_social**: Razón social registrada de la empresa.
- **primer_apellido**: Primer apellido del representante legal o persona natural (si aplica).
- **segundo_apellido**: Segundo apellido del representante legal o persona natural (si aplica).
- **primer_nombre**: Primer nombre del representante legal o persona natural (si aplica).
- **segundo_nombre**: Segundo nombre u otros nombres del representante legal o persona natural.
- **sigla**: Sigla o acrónimo de la empresa, si aplica.
- **codigo_clase_identificacion**: Código correspondiente a la clase de identificación.
- **clase_identificacion**: Descripción textual de la clase de identificación utilizada.
- **numero_identificacion**: Número de identificación del representante legal o NIT, según aplique.
- **nit**: Número de Identificación Tributaria (NIT) asignado por la DIAN.
- **digito_verificacion**: Dígito de verificación del NIT, usado para validación tributaria.
- **cod_ciiu_act_econ_pri**: Código CIIU correspondiente a la actividad económica principal.
- **cod_ciiu_act_econ_sec**: Código CIIU correspondiente a la actividad económica secundaria.
- **ciiu3**: Código CIIU para otra actividad económica registrada adicional.
- **ciiu4**: Código CIIU para una segunda actividad económica adicional.
- **fecha_matricula**: Fecha en que se registró la matrícula mercantil.
- **fecha_renovacion**: Fecha de última renovación de la matrícula.
- **ultimo_anio_renovado**: Año más reciente en que se renovó la matrícula.
- **fecha_vigencia**: Fecha de vigencia de la inscripción comercial.
- **fecha_cancelacion**: Fecha en que se canceló la matrícula, si aplica.
- **codigo_tipo_sociedad**: Código asignado al tipo societario (SAS, LTDA, etc.).
- **tipo_sociedad**: Nombre del tipo de sociedad legalmente constituida.
- **codigo_organizacion_juridica**: Código de la organización jurídica (ej. entidad sin ánimo de lucro, sociedad anónima).
- **organizacion_juridica**: Nombre de la organización jurídica de la empresa.
- **codigo_categoria_matricula**: Código que representa la categoría de la matrícula comercial.
- **categoria_matricula**: Descripción textual de la categoría de matrícula.
- **codigo_estado_matricula**: Código del estado actual de la matrícula (ej. vigente, cancelada).
- **estado_matricula**: Estado textual de la matrícula mercantil.
- **clase_identificacion_rl**: Clase de identificación del representante legal de la empresa.
- **num_identificacion_representante_legal**: Número de identificación del representante legal.
- **representante_legal**: Nombre completo del representante legal.
- **fecha_actualizacion**: Fecha de la última actualización de los datos en el RUES.
- **md5_hash**: Código hash que representa el contenido del registro para control de integridad.

Capa: **staging**

Tabla: **staging_hub_persona**

Contiene los registros centrales que identifican de forma única a cada persona en el sistema. Representa la entidad principal en el modelo Data Vault.

Campos:

- **id**: Identificador único del registro.
- **id_estadistico_persona**: Identificador estadístico asignado a la persona.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_sat_persona**

Contiene la información descriptiva y sociodemográfica de una persona como nombres, contacto, etnia, condición de discapacidad, etc.

Campos:

- **id**: Identificador único del registro.
- **id_hub_persona**: Identificador de la persona asociada (referencia al hub).
- **primer_nombre**: Primer nombre.
- **segundo_nombre**: Segundo nombre.
- **primer_apellido**: Primer apellido.
- **segundo_apellido**: Segundo apellido.
- **fecha_nacimiento**: Fecha de nacimiento.
- **telefono**: Número de teléfono.
- **email**: Dirección de correo electrónico.
- **sexo**: Sexo declarado.
- **discapacidad**: Indica si la persona presenta discapacidad (booleano).
- **tipo_poblacion_especial**: Población especial (desplazado, indígena, etc.).
- **etnia**: Grupo étnico declarado.
- **nombre_resguardo**: Nombre del resguardo indígena, si aplica.
- **estatus**: Registros temporales
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: **staging_hub_entidad**

Contiene los identificadores únicos de las entidades (por ejemplo: EPS, IPS, organizaciones, etc.) vinculadas con personas.

Campos:

- **id**: Identificador único del registro.
- **codigo_entidad**: Código único que representa a la entidad.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_sat_entidad**

Contiene información descriptiva de entidades: nombre, tipo, estado y código institucional.

Campos:

- **id**: Identificador único del registro.
- **id_hub_entidad**: Referencia al identificador de entidad en el hub.
- **nombre**: Nombre de la entidad.
- **descripcion**: Descripción general de la entidad.
- **tipo**: Tipo de entidad (pública, privada, etc.).
- **codigo_SICODE**: Código del sistema institucional.
- **estatus**: Registros temporales.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_hub_documento**

Contiene los documentos únicos utilizados para identificar a las personas.

Campos:

- **id**: Identificador único del registro.
- **numero_documento**: Número del documento de identidad.

- **load_datetime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_sat_documento

Información adicional de los documentos, como lugar y fecha de expedición.

Campos:

- **id:** Identificador único del registro.
- **id_hub_documento:** Referencia al documento registrado en el hub.
- **lugar_expedicion:** Lugar donde fue expedido el documento.
- **fecha_expedicion:** Fecha de expedición.
- **fecha_vencimiento:** Fecha de vencimiento.
- **id_tipo_documento:** Tipo de documento según catálogo.
- **load_datetime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_persona_documento

Relaciona una persona con su documento de identidad.

Campos:

- **id:** Identificador único del registro.
- **id_hub_persona:** Identificador de la persona.
- **id_hub_documento:** Identificador del documento.
- **load_datetime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_migracion_persona

Registra un evento de migración (cambio de ubicación) de una persona.

Campos:

- **id**: Identificador único del registro.
- **id_hub_persona**: Persona que realizó la migración.
- **id_ubicacion_origen**: Ubicación de origen.
- **id_ubicacion_destino**: Ubicación de destino.
- **dias_diff_sum_dentro**: Días acumulados dentro del país.
- **dias_diff_sum_fuera**: Días acumulados fuera del país.
- **dias_diff_sum_otros**: Otros días registrados.
- **fecha_migracion**: Fecha del evento de migración.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_nacimiento_persona

Registra eventos de nacimiento de personas, con relación a sus padres y ubicación.

Campos:

- **id**: Identificador único del registro.
- **id_hub_padre**: Identificador del padre.
- **id_hub_madre**: Identificador de la madre.
- **id_hub_persona**: Persona recién nacida.
- **id_ubicacion_nacimiento**: Ubicación del nacimiento.
- **fecha_nacimiento**: Fecha del nacimiento.
- **peso_nacimiento**: Peso registrado al nacer.
- **ips_primaria**: Institución prestadora de servicios de salud donde nació.
- **sexo_nacimiento**: Sexo registrado al nacer.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_fallecido_persona

Registra los eventos de fallecimiento y sus características asociadas.

Campos:

- **id**: Identificador único del registro.
- **id_hub_persona**: Persona fallecida.
- **id_ubicacion**: Ubicación del fallecimiento.
- **fecha_fallecimiento**: Fecha en que ocurrió el fallecimiento.
- **hora_fallecimiento**: Hora del fallecimiento.

- **causa_defuncion**: Causa médica del fallecimiento.
- **causa_previo_defuncion**: Condición o evento que precedió el fallecimiento.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_link_estado_civil_persona**

Registra eventos relacionados con el estado civil de las personas.

Campos:

- **id**: Identificador único del registro.
- **id_hub_persona**: Persona a la que se le registra el estado civil.
- **fecha_evento**: Fecha del cambio de estado civil.
- **descripcion_estado_civil**: Estado civil alcanzado (ej. soltero, casado).
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_hub_empresa**

Contiene la identificación única de cada empresa registrada en el sistema, como entidad central del modelo Data Vault.

Campos:

- **id**: Identificador único del registro.
- **id_estadistico_empresa**: Identificador estadístico asignado a la empresa.
- **razon_social**: Razón social de la empresa.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_sat_empresa**

Contiene información detallada y complementaria de la empresa como matrícula, tipo de sociedad, representante legal, actividad económica y vigencia legal.

Campos:

- **id**: Identificador único del registro.
- **id_hub_empresa**: Identificador de la empresa asociado (referencia a **hub_empresa**).
- **codigo_camara**: Código interno de la Cámara de Comercio que registró o actualizó la información.
- **camara_comercio**: Nombre de la Cámara de Comercio correspondiente.
- **matricula**: Número de matrícula mercantil asignado por la Cámara de Comercio.
- **inscripcion_proponente**: Número de inscripción en el Registro Único de Proponentes (RUP).
- **razon_social**: Razón social actualizada de la empresa.
- **sigla**: Sigla o acrónimo que identifica a la empresa, si aplica.
- **fecha_matricula**: Fecha de registro inicial de la matrícula mercantil.
- **fecha Renovacion**: Fecha de la última renovación de la matrícula mercantil.
- **ultimo_anio_renovado**: Último año en que se renovó la matrícula. (Debe estar entre 1999 y 2050).
- **fecha_vigencia**: Fecha hasta la cual es válida la inscripción mercantil.
- **fecha_cancelacion**: Fecha en que se canceló la matrícula, en caso de aplicar.
- **codigo_tipo_sociedad**: Código que representa el tipo de sociedad (ej. SAS, LTDA, SA).
- **tipo_sociedad**: Descripción textual del tipo de sociedad legalmente constituida.
- **codigo_organizacion_juridica**: Código del tipo de organización jurídica (ej. fundación, cooperativa, sociedad anónima).
- **organizacion_juridica**: Descripción del tipo de organización jurídica de la empresa.
- **codigo_categoria_matricula**: Código correspondiente a la categoría de la empresa en términos de tamaño o tipo.
- **categoria_matricula**: Descripción textual de la categoría empresarial (ej. micro, pequeña, mediana).
- **codigo_estado_matricula**: Código del estado del registro mercantil (activo, cancelado, suspendido).
- **estado_matricula**: Descripción del estado legal actual del registro mercantil.
- **clase_identificacion_rl**: Clase o tipo de documento del representante legal (CC, CE, NIT, etc.).
- **num_identificacion_representante_legal**: Número de identificación del representante legal.
- **nombre_representante_legal**: Nombre completo del representante legal registrado.
- **fecha_actualizacion**: Fecha en la que se actualizó la información más recientemente en el RUES.
- **codigo_clase_identificacion**: Código del tipo de identificación fiscal asignado a la empresa.

- **numero_identificacion_tributario**: Número de Identificación Tributaria (NIT) asignado por la DIAN.
- **digito_verificacion**: Dígito de verificación del NIT, usado para validar consistencia.
- **estatus**: Registros temporales
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_representante_legal_empresa

Establece la relación entre una empresa y su representante legal.

Campos:

- **id**: Identificador único del registro.
- **id_hub_representante_legal**: Identificador de la persona que actúa como representante legal.
- **id_hub_empresa**: Identificador de la empresa representada.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: staging_link_empresa_actividad_economica

Relaciona una empresa con una o varias actividades económicas, utilizando códigos CIIU.

Campos:

- **id**: Identificador único del registro.
- **id_hub_empresa**: Empresa a la que se vincula la actividad.
- **id_ciiu**: Código CIIU correspondiente a la actividad.
- **es_primario**: Indica si la actividad es la principal (**true**) o secundaria (**false**).
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Tabla: **staging_link_persona_afiliacion**

Relaciona a una persona afiliada con su cotizante, aportante (empresa), lugar de afiliación y estado dentro del sistema de salud.

Campos:

- **id**: Identificador único del registro.
- **id_hub_persona_afiliada**: Identificador de la persona afiliada.
- **id_hub_documento_afiliado**: Documento de identidad del afiliado.
- **id_hub_persona_cotizante**: Identificador de la persona que cotiza.
- **id_hub_documento_cotizante**: Documento del cotizante.
- **id_tipo_documento_cotizante**: Tipo de documento del cotizante.
- **id_hub_aportante**: Empresa o entidad que realiza el aporte (hub_empresa).
- **id_hub_entidad**: Identificador de la entidad.
- **id_tipo_documento_aportante**: Tipo de documento de la entidad aportante.
- **id_cod_departamento**: Departamento donde se realizó la afiliación.
- **id_cod_municipio**: Municipio donde se realizó la afiliación.
- **id_zona**: Zona geográfica (urbana, rural) desde donde se afilió.
- **tipo_cotizante**: Tipo de cotizante (dependiente, independiente, etc.).
- **tipo_afiliado**: Clasificación del afiliado (cotizante, beneficiario, etc.).
- **parentesco**: Relación del afiliado con el cotizante.
- **ips_primaria**: Institución prestadora de servicios de salud asignada.
- **fecha_afiliacion**: Fecha en que se efectuó la afiliación.
- **regimen**: Régimen de afiliación (contributivo, subsidiado, especial, etc.).
- **estado_afiliacion**: Estado actual del afiliado (activo, suspendido, etc.).
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.

Capa: **bi**

Tabla: **bi_dim_tiempo**

Contiene atributos derivados del calendario asociados a fechas utilizadas en los hechos del sistema.

Campos:

- **id**: Identificador numérico único de la fecha (ej. 20240422).
- **fecha**: Fecha en formato calendario.

- **semestre:** Semestre al que pertenece la fecha (1 o 2).
- **trimestre:** Trimestre del año.
- **cuatrimestre:** Cuatrimestre del año.
- **dia:** Día del mes.
dia_del_anio: Día del año (1 a 365/366).
- **dia_semana:** Día de la semana (1 = lunes).
- **mes:** Mes del año (1-12).
- **nombre_dia:** Nombre textual del día (ej. lunes).
- **nombre_mes:** Nombre textual del mes (ej. abril).
- **abreviacion_dia:** Abreviación del día (ej. lun).
- **abreviacion_mes:** Abreviación del mes (ej. abr).

Tabla: **bi_dim_ubicacion_geografica**

Contiene información geográfica jerárquica (departamento, municipio y dirección) utilizada en los hechos de localización.

Campos:

- **id:** Identificador único de la ubicación.
- **codigo_departamento:** Código DIVIPOLA del departamento.
- **nombre_departamento:** Nombre del departamento.
- **latitud_departamento:** Latitud del centroide del departamento.
- **longitud_departamento:** Longitud del centroide del departamento.
- **nombre_municipio:** Nombre del municipio.
- **codigo_municipio:** Código DIVIPOLA del municipio.
- **latitud_municipio:** Latitud del municipio.
- **longitud_municipio:** Longitud del municipio.
- **tipo_zona:** Clasificación de la zona (urbana, rural, etc.).
- **llave_direccion:** Clave única de la dirección.
- **direccion:** Dirección física o texto de ubicación.
- **latitud:** Latitud de la ubicación específica.
- **longitud:** Longitud de la ubicación específica.

Tabla: **bi_dim_persona**

Contiene atributos demográficos y personales para análisis a nivel de individuo.

Campos:

- **id:** Identificador único de la dimensión persona.
- **id_estadistico_persona:** Identificador estadístico unificado de la persona.
- **primer_nombre:** Primer nombre.
- **segundo_nombre:** Segundo nombre.
- **primer_apellido:** Primer apellido.
- **segundo_apellido:** Segundo apellido.

- **fecha_nacimiento:** Fecha de nacimiento.
- **telefono:** Número telefónico de contacto.
- **email:** Dirección de correo electrónico.
- **sexo:** Sexo registrado.
- **discapacidad:** Indica si tiene alguna discapacidad (booleano).
- **tipo_poblacion_especial:** Clasificación como población especial.
- **etnia:** Etnia registrada.
- **nombre_resguardo:** Nombre del resguardo indígena (si aplica).
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.

Tabla: bi_dim_documento

Contiene la información de documentos de identificación relacionados con personas en los hechos.

Campos:

- **id:** Identificador único del documento.
- **abreviacion_documento:** Abreviación del tipo de documento (ej. CC, TI).
- **nombre_documento:** Nombre completo del documento.
- **numero_documento:** Número del documento.
- **tipo_documento:** Tipo de documento (descripción).
- **descripcion:** Descripción extendida o anotación adicional.
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.

Tabla: bi_dim_estado_civil

Contiene los estados civiles disponibles en el sistema, usados en hechos personales.

Campos:

- **id:** Identificador único del estado civil.
- **descripcion_estado_civil:** Estado civil registrado (ej. soltero, casado).
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.

- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: bi_fact_afiliacion

Hecho que registra relaciones de afiliación entre personas (cotizante–afiliado) y condiciones asociadas.

Campos:

- **id**: Identificador único del hecho.
- **id_persona_cotizante**: Referencia a la persona cotizante.
- **id_persona_afiliada**: Referencia a la persona afiliada.
- **id_documento_cotizante**: Documento del cotizante.
- **id_documento_afiliado**: Documento del afiliado.
- **parentesco**: Relación entre cotizante y afiliado.
- **ips_primaria**: Institución de salud asignada.
- **fecha_afiliacion**: Fecha de la afiliación (referencia a dim_tiempo).
- **regimen**: Régimen de afiliación (contributivo, subsidiado, etc.).
- **estado_afiliacion**: Estado actual del afiliado.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: bi_fact_migracion_persona

Hecho que registra movimientos de una persona entre diferentes ubicaciones.

Campos:

- **id**: Identificador único del hecho.
- **id_dim_persona**: Persona que realiza la migración.
- **id_dim_ubicacion_origen**: Ubicación de origen.
- **id_dim_ubicacion_destino**: Ubicación de destino.
- **id_dim_fecha_migracion**: Fecha del evento migratorio (referencia a dim_tiempo).
- **dias_diff_sum_dentro**: Días acumulados dentro del país.
- **dias_diff_sum_fuera**: Días acumulados fuera del país.
- **dias_diff_sum_otros**: Días sin clasificar o en tránsito.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: **bi_fact_fallecimiento_persona**

Hecho que registra el fallecimiento de una persona y sus características.

Campos:

- **id**: Identificador único del hecho.
- **id_dim_persona**: Persona fallecida.
- **id_dim_hora_fallecimiento**: Fecha/hora de fallecimiento (referencia a dim_tiempo).
- **id_dim_ubicacion_fallecimiento**: Ubicación donde ocurrió el fallecimiento.
- **id_fecha_fallecimiento**: Fecha del fallecimiento (referencia a dim_tiempo).
- **causa_defuncion**: Causa médica del fallecimiento.
- **causa_previo_defuncion**: Antecedente que provocó la defunción.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: **bi_fact_nacimiento_persona**

Hecho que registra un evento de nacimiento, incluyendo padres, ubicación y peso.

Campos:

- **id**: Identificador único del hecho.
- **id_dim_persona**: Persona recién nacida.
- **id_dim_padre**: Padre.
- **id_dim_madre**: Madre.
- **id_dim_ubicacion_nacimiento**: Lugar del nacimiento.
- **id_dim_fecha_nacimiento**: Fecha del nacimiento.
- **peso_nacimiento**: Peso en kilogramos.
- **sexo_nacimiento**: Sexo registrado al nacer.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: **bi_fact_estado_civil_persona**

Hecho que registra eventos de cambio de estado civil en una persona.

Campos:

- **id**: Identificador único del hecho.
- **id_dim_persona**: Persona relacionada.

- **id_dim_estado_civil**: Estado civil alcanzado.
- **id_dim_fecha_evento**: Fecha del cambio (referencia a dim_tiempo).
- **observacion**: Observación adicional del evento.
- **md5_hash**: Valor hash calculado para control de integridad y detección de duplicados.
- **load_datatime**: Fecha en que se registró el dato en el sistema.
- **id_fuente_registro**: Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl**: Referencia al proceso ETL que realizó la carga del registro.

Tabla: **bi_dim_empresa**

Contiene la dimensión descriptiva de empresas. Se usa para análisis históricos y cruzados con hechos relacionados a actividades económicas, registros, y estructuras legales. Esta dimensión puede estar parcialmente desnormalizada para eficiencia analítica.

Campos:

- **id**: Identificador único del registro en la dimensión.
- **id_estadistico_empresa**: Identificador estadístico único de la empresa.
- **codigo_camara**: Código de la Cámara de Comercio que realiza la actualización de datos.
- **camara_comercio**: Nombre de la Cámara de Comercio correspondiente.
- **matricula**: Número de Matrícula Mercantil asignada por la Cámara de Comercio.
- **inscripcion_proponente**: Número de inscripción en el Registro Único de Proponentes.
- **razon_social**: Razón social registrada de la empresa.
- **sigla**: Sigla o acrónimo de la empresa, si aplica.
- **fecha_matricula**: Fecha en que se registró la matrícula mercantil.
- **fecha_renovacion**: Fecha de la última renovación de la matrícula.
- **ultimo_anio_renovado**: Año más reciente en que se renovó la matrícula (validación entre 1999 y 2050).
- **fecha_vigencia**: Fecha de vigencia de la inscripción comercial.
- **fecha_cancelacion**: Fecha en que se canceló la matrícula, si aplica.
- **codigo_tipo_sociedad**: Código asignado al tipo societario (SAS, LTDA, SA, etc.).
- **tipo_sociedad**: Nombre del tipo de sociedad legalmente constituida.
- **codigo_organizacion_juridica**: Código de la organización jurídica (ej. entidad sin ánimo de lucro, sociedad anónima).
- **organizacion_juridica**: Nombre de la organización jurídica de la empresa.
- **codigo_categoria_matricula**: Código que representa la categoría de la matrícula comercial.
- **categoria_matricula**: Descripción textual de la categoría de matrícula (ej. micro, pequeña, mediana).
- **codigo_estado_matricula**: Código del estado actual de la matrícula (ej. vigente, cancelada).

- **estado_matricula:** Estado textual de la matrícula mercantil.
- **clase_identificacion_rl:** Clase de identificación del representante legal de la empresa (CC, CE, NIT, etc.).
- **num_identificacion_representante_legal:** Número de identificación del representante legal.
- **nombre_representante_legal:** Nombre completo del representante legal.
- **fecha_actualizacion:** Fecha de la última actualización de los datos en el RUES.
- **codigo_clase_identificacion:** Código correspondiente a la clase de identificación fiscal.
- **numero_identificacion_tributario:** Número de Identificación Tributaria (NIT) asignado por la DIAN.
- **digito_verificacion:** Dígito de verificación del NIT, usado para validación tributaria.
- **codigo_ciiu_principal :** Código CIIU principal de actividad económica.
- **actividad_economica_principal:** Descripción de la actividad económica principal.
- **codigo_ciiu_secundaria:** Código CIIU secundario.
- **actividad_economica_secundaria:** Descripción de la actividad económica secundaria.
- **codigo_ciiu_terciario:** Código CIIU terciario
- **actividad_economica_ciiu3:** Descripción de la tercera actividad económica registrada
- **codigo_ciiu_cuaternario:** Código CIIU cuaternario
- **actividad_economica_ciiu4:** Descripción de la cuarta actividad económica registrada.
- **md5_hash:** Valor hash calculado para control de integridad y detección de duplicados.
- **load_datetime:** Fecha en que se registró el dato en el sistema.
- **id_fuente_registro:** Identificador de la fuente oficial que proporcionó el dato.
- **id_log_etl:** Referencia al proceso ETL que realizó la carga del registro.