

Información sobre la BD_Autos1

Agencia

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Agencia

Estas son las entidades que se encuentran en la BD_Autos1.

La entidad Agencia está relacionada con Dirección y Auto, ya que cada Agencia tiene una Dirección, y los autos son vendidos en agencias.

¿Cómo se define cada entidad?

Esta entidad se define Agencia, la cual es la representación de un local en donde se venden autos.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- RFC: Es el identificador de la agencia y es único. No se permiten nulos.
- NOMBRE: Es el nombre al que se le pone a la Agencia.
- CORREO: Es el correo electrónico en el cual se puede contactar a la Agencia.
- TELEFONO: Es el número telefónico en el cual puedes contactar a la Agencia.
- DIRECCION: Es la dirección en la cual se encuentra ubicada.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- RFC: VARCHAR
- NOMBRE: VARCHAR
- CORREO: VARCHAR
- TELEFONO: INTEGER
- DIRECCION: INTEGER

Identificar llaves primarias.

El RFC es la llave primaria.

Información de la procedencia de los datos.

La información está guardada en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la práctica 3.

Auto

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Auto

La entidad auto se relaciona con Operación y Realiza, ya que el no_motor lo lleva la realiza y pre_compra tiene el mismo atributo.

¿Cómo se define cada entidad?

La entidad Auto es la representación de un automóvil, el cual describe a la perfección todas sus características.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- no_motor: es el identificador del vehículo, ya que este número identifica a el vehículo. Su dominio debe ser cualquier número mayor a 0 y debe ser único.
- hp: caballos de fuerza que tiene el vehículo. Su dominio debe ser cualquier número mayor a 0.
- anio: año en el cual fue lanzado el vehículo. Su dominio debe ser cualquier año.
- km: número de kilómetros que tiene el vehículo. Su dominio debe ser cualquier número mayor a 0.
- marca: Marca a la cual pertenece el vehículo. Su dominio es cualquier marca de autos.
- modelo: Tipo de modelo del vehículo. Su dominio depende de la marca de autos, y debe de ser algún modelo de la marca.
- versión: Tipo de versión del vehículo. Su dominio depende del modelo del auto, y debe ser alguna versión del modelo.
- color: Tipo de color del vehículo. Su dominio es cualquier color.
- categoría: Decide la categoría que pertenece el vehículo. Su dominio debe ser AUTO o CAMIONETA.
- tipo: Tipo de vehículo. Su dominio debe ser AUTO o CAMIONETA.
- km_l: cuantos kilómetros rinde un litro. Su dominio es cualquier número decimal mayor a 0.
- transmisión: que tipo de transmisión tiene el vehículo. Su dominio es AUTOMATED MANUAL-SELECTABLE, AUTOMATEDCMANUAL, CONTINUOUSLY VARIABLE, MANUAL, OTHER, AUTOMATIC, SEMI-AUTOMATIC o SELECTABLE CONTINUOUSLY VARIABLE.
- no_puertas: número de puertas que tiene el vehículo. Su dominio es cualquier número mayor a 3.
- nuevo: nos dice si el carro es nuevo o usado. Su dominio son los valores TRUE o FALSE

- no_cilindros: cuantos cilindros tiene el vehículo. Su dominio es cualquier numero mayor a 4.
- fecha_compra: cuando fue comprado el vehículo. Su dominio es cualquier fecha del calendario, pero debe ser menor al día actual.
- pr_compra: precio el cual deberá ser comprado el auto. Su dominio es cualquier valor de tipo money mayor a 0.
- rfc_agencia: identificador de la agencia en la cual el vehículo esta en exhibición. Su dominio dependerá de el numero de agencias que hallamos registrado.

¿Cuál es la representación de valores ausentes?
NULL

Identificar los tipos de datos de cada atributo.

- no_motor: INTEGER
- hp: INTEGER
- anio: INTEGER
- km: INTEGER
- marca: VARCHAR
- modelo: VARCHAR
- versión: VARCHAR
- color: VARCHAR
- categoría: VARCHAR
- tipo: VARCHAR
- km_l: REAL
- transmisión: VARCHAR
- no_puertas: INTEGER
- nuevo: BOOLEAN
- no_cilindros: INTEGER
- fecha_compra: DATE
- pr_compra: MONEY
- rfc_agencia: VARCHAR

Identificar llaves primarias.
La llave primaria es no_motor

Información de la procedencia de los datos.

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Cliente

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Cliente

La relación Cliente esta relacionado con Persona y Realiza, ya que cada cliente es una Persona y los clientes Realizan compras de vehículos.

¿Cómo se define cada entidad?

La entidad Cliente identifica a una persona como cliente, la cual nos dice su grado de estudio y los ingresos mensuales.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- rfc_cliente: es la clave federal de cada persona y es única. Su dominio depende de cuantas personas fueran de altas, y a una persona solo se le asigna un único valor.
- Ingresos: la cantidad de dinero que tiene de ingresos esa persona. Su dominio es cualquier valor money mayor a 0.
- grado_estudio: el grado de estudios que tiene la persona. Su dominio es: PRIMARIA, SECUNDARIA, PREPARATORIA, UIVERSIDAD, MAESTRIA o DOCTORADO.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- rfc_clienter: VARCHAR
- ingresos: MONEY
- grado_estudio: VARHCAR

Identificar llaves primarias.

La llave primaria es el rfc_cliente

Información de la procedencia de los datos.

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Dirección

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Dirección

¿Cómo se define cada entidad?

- Esta entidad gestiona la ubicación, a fin de tener una referencia a entidades existentes

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- Id: es el que identifica la dirección. Su dominio es un valor numérico.
- Calle: es el nombre de la calle donde se reside. Su dominio es cualquier nombre registrado como calle.
- Colonia: es el nombre de la colonia donde se reside. Su dominio es cualquier nombre registrado como colonia.
- Delegación_municipio: es el estado o delegación donde se reside. Su dominio es cualquier nombre registrado como colonia.
- Estado: Estado en el que se reside. Su dominio es cualquier estado del país.
- Cp: es la clave la cual identifica la colonia. Su dominio es un número diferente de 0.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- Id: INTEGER
- Calle: VARCHAR
- Colonia: VARCHAR
- Delegación_municipio: VARCHAR
- Estado: VARCHAR
- Cp: INTEGER

Identificar llaves primarias.

LA llave primaria es id.

Información de la procedencia de los datos.

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Operación

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Operación

La entidad Operación está relacionada con Auto, ya que nos dice cuanto cuesta el auto.

¿Cómo se define cada entidad?

Esta entidad representa los movimientos que se generan en la agencia.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- Id_operacion:
- Método_pago
- Fecha_venta
- Pr_venta
- Descuento
- Pr_venta_final

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- Id_operacion: INTEGER
- Método_pago: VARCHAR
- Fecha_venta: DATE
- Pr_venta: MONEY
- Descuento: VARCHAR
- Pr_venta_final: MONEY

Identificar llaves primarias.

La llave primaria id_operacion.

Información de la procedencia de los datos.

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Persona

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Persona

No tiene ninguna entidad relacionada.

¿Cómo se define cada entidad?

Es la entidad que nos da la información sobre las personas

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- Rfc: es la clave federal de cada persona y es única. Su dominio depende de cuantas personas fueran de altas, y a una persona solo se le asigna un único valor.
- Nombre: Nombre de la persona. Su dominio es cualquier nombre.
- Apaterno: primero apellido de la persona. Su dominio es cualquier apellido.
- Amaterno: segundo apellido de la persona. Su dominio es cualquier apellido.
- Correo: correo electrónico de la persona. Su dominio es cualquier dirección de correo.
- Genero: define su genero. Su dominio es M o F.
- Teléfono: número de teléfono de la persona. Su dominio es una cadena de caracteres con mínimo de 10 dígitos.
- Anio_nacimiento: año de nacimiento de la persona. Su dominio es cualquier número mayor a 1900.
- Id_direccion: identificador de la dirección donde reside la persona. Su dominio es cualquier número existente id_direccion en la entidad Dirección.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- Rfc: VARCHAR
- Nombre: VARCHAR
- Apaterno: VARCHAR
- Amaterno: VARCHAR
- Correo : VARCHAR
- Genero: CHAR
- Teléfono: INTEGER
- Anio_nacimiento: INTEGER
- Id_direccion: INTEGER

Identificar llaves primarias.

La llave primaria es rfc.

Información de la procedencia de los datos

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Realiza

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Realiza

La entidad Realiza esta relacionado con Vendedor, Cliente, Auto y Operación, ya que nos da el rfc del vendedor y del cliente, nos dice que operación se realizo y que auto fue involucrado.

¿Cómo se define cada entidad?

Esta entidad nos dice que tipo de operación fue realizada, que vendedor la realizo y que cliente estuvo involucrado.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- Rfc_vendedor: es la clave federal de cada vendedor y es única. Su dominio depende de cuantas personas fueran de altas, y a una persona solo se le asigna un único valor.
- Rfc_cliente: es la clave federal de cada cliente y es única. Su dominio depende de cuantas personas fueran de altas, y a una persona solo se le asigna un único valor.
- No_motor: es el identificador del vehículo, ya que este número identifica a el vehículo. Su dominio debe ser cualquier número mayor a 0 y debe ser único.
- Id_operacion: identificador de la operación. Su dominio es cualquier número existente id_operacion que este en la entidad Operación

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- Rfc_vendedor: VARCHAR
- Rfc_cliente: VARCHAR
- No_motor: VARCHAR
- Id_operacion: INTEGER

Identificar llaves primarias.

Es una llave compuesta de estos atributos rfc_vendedor, rfc_cliente, no_motor, id_operacion.

Información de la procedencia de los datos

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Vendedor

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- Vendedor

Esta entidad se relaciona con Agencia, Realiza y Persona. Ya que cada vendedor trabaja en una agencia, los vendedores realizan las ventas y los vendedores son personas.

¿Cómo se define cada entidad?

La entidad Vendedor es la que identifica a una persona como vendedor, el cual nos dice cuanto gana mensualmente y en que agencia trabaja.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- rfc_vendedor: Identificador de la persona que es vendedor. Su dominio depende de cuantas personas fueron registradas.
- sueldo: es la paga que se le hace al vendedor. Su dominio es cualquier numero mayor a 0.
- rfc_agencia: Identificador de las agencias en la que trabaja.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- rfc_vendedor: VARCHAR
- sueldo: MONEY
- rfc_agencia: VARCHAR

Identificar llaves primarias.

No tiene llave primaria.

Información de la procedencia de los datos

La información esta guardado en una base de datos llamada BD_Autos1, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Información sobre la base de datos BD_Autos2

¿Qué entidades posee el conjunto de datos? y ¿cómo se relacionan?

- **Dirección:** Esta entidad se relaciona con las entidades “Persona”, “Vendedor” y “Agencia”, con una cardinalidad de 1:1 en ambas entidades. Ya que tanto una agencia, como cualquier persona debe tener una y solo una dirección.
- **Persona:** En este caso, la entidad Persona es una superentidad con todos los atributos, donde las subentidades que son “Vendedor” y “Cliente” contienen el atributo llave de la entidad superior.
- **Vendedor:** Es una subentidad de la entidad Persona en la que se relaciona con la entidad “Operación” con una cardinalidad de 1:n
- **Cliente:** Al igual que la entidad Vendedor, Cliente es una subentidad de la entidad Persona y se relaciona con las entidades Operación y Auto
- **Agencia:** Esta entidad está relacionada con la entidad Auto y Direccion. Con la entidad Auto esta relacionado con una cardinalidad de 1:n ya que una agencia puede adquirir varios automóviles
- **Auto:** Esta relacionada con las entidades “Agencia”, “Vendedor”, “Operación” y “Cliente”. Este tipo de relación es conocida como N-aria.
- **Operación:** Al igual que la entidad “Auto” esta entidad cumple con las mismas propiedades ya antes definidas.

¿Cómo se define cada entidad?

- **Dirección:** Esta entidad gestiona la ubicación, a fin de tener una referencia a entidades existentes
- **Persona:** Esta entidad representa tanto a los clientes como trabajadores de una agencia de automóviles.
- **Vendedor:** La entidad Vendedor es la que identifica a una persona como vendedor, el cual nos dice cuanto gana mensualmente y en que agencia trabaja.

- **Cliente:** Al igual que la entidad Vendedor, Cliente es una entidad que identifica a una persona como cliente
- **Agencia:** Esta entidad representa a una institución que se encarga de comercializar automóviles.
- **Auto:** Representa el producto, junto con todas sus características que venden las agencias.
- **Operación:** Al igual que la entidad "Auto" esta entidad lleva el registro de las diferentes operaciones que se llevan a cabo en la agencia.

¿Cómo se define cada atributo? y ¿cuál es su dominio de datos (valores válidos)?

- **Fecha compra:** Es la fecha en que fue realizada la compra del automóvil
- **Modelo:** Tipo de modelo del vehículo, donde su dominio debe ser algún modelo de la marca
- **Version:** Tipo de versión del vehículo, donde su dominio debe ser alguna versión del modelo
- **Hp:** Caballos de fuerza que tiene el vehículo. Su dominio debe ser cualquier número mayor a 0.
- **max Velocidad:** Es la velocidad máxima que el automóvil puede alcanzar
- **0 - 100:** Es el tiempo que le toma al automóvil de ir desde velocidad cero hasta 100, su dominio debe ser cualquier valor mayor a 0.
- **1/100km:** Es la distancia que puede recorrer el vehículo. Su dominio es cualquier valor numérico positivo.
- **Transmisión:** Es el tipo de transmisión que tiene el automóvil. Su dominio es la variación que hay entre los automóviles Manuales y Automáticos.
- **Categoría:** Son los tipos de categoría que pertenece el vehículo. Su dominio debe ser Deportivo o Camioneta.
- **Color:** Es el tipo de color del vehículo. Su dominio puede ser cualquier color.
- **Accesorios:** Son los tipos de accesorios que el vehículo puede tener. Su dominio pueden ser quemacocos, asientos de lujo o cualquier otro objeto que complementa el vehículo.
- **Costo total compra:** Es la cifra total de la compra. Su dominio debe ser cualquier valor numérico mayor a 0.
- **Precio de venta:** Es el costo del producto. Su dominio debe ser cualquier valor numérico mayor a 0.
- **Enganche:** Es una proporción del valor del inmueble que depende de un porcentaje. Su dominio es cualquier porcentaje excepto, negativos.
- **Curp:** Es la clave federal de cada persona y es única. Su dominio depende de cuantas personas fueron dadas de alta donde se le asigna un único valor.
- **Nombre cliente:** Es el nombre de cada cliente que se registra en la base. Su dominio puede ser cualquier carácter excepto valores ASCII.

- **Edad:** Es la edad de cada cliente registrado en la base. Su dominio debe ser cualquier valor numérico mayor a 0.
- **Genero cliente:** Es el genero del cliente. Su dominio son Masculino, Femenino.
- **Ingresos:** Es el campo que lleva un registro de los ingresos del cliente. Su dominio es cualquier valor numérico.
- **Grados Estudio:** Es el ultimo grado de estudios del cliente. Su dominio puede ser cualquier grado académico.
- **Dirección:** Es la dirección del cliente registrado en la base de datos. Su dominio es el Código Postal y el estado donde se vive.
- **Teléfono:** Número de teléfono del cliente. Su dominio es un entero con mínimo de 10 dígitos.
- **Correo electrónico:** Correo electrónico del cliente. Su dominio puede ser cualquier extension de correo valido.
- **Fecha venta:** Fecha en que se realizó la venta. Su dominio es cualquier fecha con formato de dd/mm/nn
- **Clave de vendedor:** Clave única del vendedor. Su dominio puede ser cualquier valor numérico.
- **Sueldo:** Sueldo de los vendedores que trabajan en las agencias. Su dominio es cualquier valor numérico mayor a 0.
- **Comisión:** Es la cantidad que se cobra por realizar transacciones que corresponden a un porcentaje. Su dominio es cualquier entero mayor a 0.
- **Nombre del Vendedor:** Nombre y apellidos del vendedor. Su dominio es cualquier nombre.
- **Genero del Vendedor:** Es el genero del cliente. Su dominio son Masculino y Femenino
- **Agencia:** Es el nombre de la agencia. Su dominio puede ser cualquier cadena.
- **Ubicación de agencia:** Dirección donde esta situada la agencia. Su dominio son Código Postal donde esta ubicado.
- **Teléfono agencia:** Numero de teléfono de la agencia. Su dominio puede ser cualquier entero con un mínimo de 10 valores.

¿Cuál es la representación de valores ausentes?

NULL

Identificar los tipos de datos de cada atributo.

- **Fecha compra:** String
- **Modelo:** String
- **Version:** String
- **Hp:** Integer
- **max Velocidad:** Integer
- **0 - 100:** Number
- **1/100km:** Number
- **Transmisión:** String
- **Categoría:** String
- **Color:** String
- **Accesorios:** String
- **Costo total compra:** Integer

- **Precio de venta:** Integer
- **Enganche:** String
- **Curp:** String
- **Nombre cliente:** String
- **Edad:** Integer
- **Genero cliente:** String
- **Ingresos:** Number
- **Grados Estudio:** String
- **Dirección:** String
- **Telefono:** Integer
- **Correo electrónico:** String
- **Fecha venta:** String
- **Clave de vendedor:** Integer
- **Sueldo:** Number
- **Comisión:** Integer
- **Nombre del Vendedor:** String
- **Genero del Vendedor:** String
- **Agencia:** String
- **Ubicación de agencia:** String
- **Telefono agencia:** String

Identificar llaves primarias.

- Las Llaves primarias son:
- Para cliente su llave primaria es: CURP.
- Para vendedor su llave primaria es: Clave vendedor.

Identificar el nivel de granularidad:

- Datos de ventas o compras de un vehículo, pueden registrarse por dia, semana, mes, trimestre, semestre, anual.

Identificar si existen atributos que:

- Son parte de un mismo atributo y se fraccionaron
 - Para el caso del atributo “nombre” se fracciono a los atributos de “nombre”, “Apaterno” y “Amaterno”. Al igual que en el atributo “dirección” se creo una entidad y se fracciono sus atributos a “Calle”, “Numero”, “C.P.”, “Estado”, etc.
- Están fraccionados y deben unificarse.
 - Para el caso de la entidad “Auto” varios de sus atributos tuvieron que unificarse ya que varios de ellos no coincidían entre tablas por lo que el conjunto integrado se asegurara de que en ambos casos los datos vayan al lugar correcto.

Información de la procedencia de los datos.

La información esta guardada en un archivo cvs el cual lleva el nombre de BD_Autos2, la cual fue proporcionada por el ayudante de laboratorio de Almacén y Minería de Datos, la cual forma parte de la practica 3.

Unificación:

Para el proceso de unificación fue necesario el acceso a los datos, en este caso los archivos cvs.

Una vez localizado los datos procedimos con la integración de modo que los registros de una fuente de datos mapeen registros en otra.

Para esto analizamos las dos fuentes de datos verificando que atributos no coincidían entre ellas y hacer una selección de las que si lo hacían.

Como resultado de la integración ahora se pueden aprovechar los diversos tipos de datos, así como mejorar la toma de decisiones y eliminar el doble trabajo.

El resultado de este proceso se puede observar en los documentos modelo/er.png y en modelo/relacional.png.