

**PROYECTO INGENIERIA DE DATOS**

**EMPRESA INVERTRANS**

**INTEGRANTES:**

Juan Nicolas Arévalo Lemus

William Alejandro Ramírez Montaña

Juan Sebastián Arroyo Zúñiga

**Bogotá D.C**

**25/02/2025**

**TABLA DE CONTENIDOS**

**Resumen Ejecutivo 3**

**Necesidad y contextualización de la necesidad 3**

**Definición del problema 3**

**Objetivos 4**

Objetivo General 4

Objetivos Específicos 4

**Alcance de la aplicación**

**Alternativas de Solución**

Soluciones existentes

Posibles soluciones:

* Solución por Juan Arroyo
* Solución por Juan Arévalo
* Solución por William Ramirez

Alternativa de solución seleccionada

**Limitaciones**

**Metodología**

**RQF’s**

**Diagramas de clases**

**Mapa StakeHolders**

**Diagrama de clases**

**Modelo conceptual**

**Modelo Lógico**

**Modelo Físico**

**Diccionario de datos**

**Scripts DDL y DML en MySQL**

**Referencias**

**RESUMEN EJECUTIVO**

*Somos estudiantes de MACC e Ingeniería de Sistemas de la Universidad del Rosario. Nuestro proyecto de semestre consiste en crear una base de datos relacional funcional, así como una no relacional. El cliente necesita esto específicamente porque tiene una compañía de transporte y hasta el día de hoy maneja toda la información relevante (facturas, conductores, historial de viajes) en un conjunto de hojas de Excel, lo cual es muy ineficiente ya que la información esta guardada en etiquetas no estandarizadas que hacen que sea muy difícil almacenar y acceder a la información de manera rápida y eficiente. Todo esto es muy perjudicial para el buen manejo de una empresa.*

*Por lo que planeamos hacer las bases de datos -relacional y no relacional como alternativa- desde cero con los conceptos aprendidos durante todo el semestre en la clase de ingeniería de datos. Nuestro objetivo es brindarle al cliente una base de datos funcional al final del semestre para que reemplace el anterior sistema de manejo de información dentro de la empresa y mejore el funcionamiento de esta. La información que manejaremos será la misma de las hojas de Excel, haciendo un ETL sobre esta para quitar variables repetidas y hacer la carga más manejable. Esto combinado con un workshop -un espacio de diálogo entre los ingenieros y el cliente en donde el cliente comparte que es lo que necesita y espera de la futura base de datos- son la base del trabajo. El proyecto será monitoreado por los propios miembros y la profesora haciendo uso de metodologías agiles como Kanban para mejorar la carga de trabajo y la eficiencia.*

**NECESIDAD Y CONTEXTUALIZACION DE LA NECESIDAD**

Nuestro cliente posee una empresa de transporte, y actualmente almacena toda la información como gastos, ganancias, nombre de clientes, conductores y placa de los vehículos en un Excel lo cual es un problema ya que no permite realizar comparaciones que determinarían decisiones como escoger entre uno u otro cliente, despedir conductores etc. Además, las búsquedas en Excel no son precisas ya que el registro de información da lugar a duplicados que significan lo mismo, un ejemplo de esto es la variación de la palabra “gastos parqueadero” que también se puede encontrar como “gasto parqueadero” “gastos parqueo” o simplemente “parqueo”. Lo que queremos brindarle a nuestro cliente es una base de datos que le permita realizar consultas y comparaciones con datos limpios para que así pueda generar reportes que le ayudaran a tomar decisiones clave.

**DEFINICIÓN DEL PROBLEMA**

Con esta en cuenta, el equipo de trabajo formuló la siguiente pregunta problema ¿Cómo podemos transformar la base de datos actual del cliente, construida en hojas de Excel con datos duplicados y etiquetas no estandarizadas, en un sistema eficiente y confiable que permita realizar consultas claras y tomar decisiones estratégicas a partir de datos limpios y organizados?

**OBJETIVOS**

**OBJETIVO GENERAL**

Construir una base de datos que facilite la comparación y consulta de información clave de una empresa de transporte, como viajes más rentables, vehículos con mayor cantidad de reparaciones o conductores con mayor número de infracciones de transito

**OBJETIVOS ESPECIFICOS**

* Construir el diagrama de clases para la base de datos de una empresa de transporte
* Proporcionar una base de datos limpia y organizada que centralice la información de la empresa de transporte
* Garantizar la capacidad de modificación y actualización de la base de datos
* Habilitar la consulta de facturas por datos esenciales como ganancias o cliente.
* Facilitar la consulta de conductores por cedula de ciudadanía
* Brindar la capacidad de consulta de vehículos por placa
* Autorizar la consulta de clientes por Id
* Conceder la consulta de conductores por número de multas
* Permitir consultar el valor promedio de un gasto por Id

**ALCANCE DE LA APLICACIÓN**

El presente proyecto tiene como objetivo diseñar e implementar una base de datos relacional y una base de datos no relacional para una empresa de transporte que actualmente gestiona su información en hojas de cálculo de Excel. A lo largo del semestre, se desarrollarán distintos módulos de trabajo, siguiendo metodologías ágiles (Kanban), y estimando los tiempos con base en **puntos de esfuerzo**, asignados conforme a la complejidad de cada tarea. A continuación, está la tabla mostrando lo que se pretende realizar en cada módulo a desarrollar:

|  |  |  |
| --- | --- | --- |
| **Módulo** | **Descripción** | **Puntos de Esfuerzo** |
| Workshop inicial | Reunión con el cliente para identificar necesidades, recopilar requerimientos y establecer criterios de éxito. | 8 (1 – 1.5 días) |
| Proceso ETL | Extracción de datos desde archivos Excel, limpieza de duplicados, estandarización de etiquetas y carga de datos estructurados. | 5 (1 día) |
| Modelado de base relacional | Diseño del diagrama entidad-relación, implementación en MySQL, inserción y validación de datos. | 13 (2 – 3 días) |
| Modelado de base no relacional | Diseño de estructura en MongoDB con documentos JSON, definición de colecciones y relaciones anidadas. | 13 (2 – 3 días) |
| Implementación de consultas | Desarrollo de consultas clave, triggers y procedimientos para facturación, vehículos, conductores, clientes y gastos promedio, tanto en SQL como en MongoDB. | 13 (2 – 3 días) |
| Pruebas y validación | Revisión funcional con el cliente, verificación de resultados esperados y ajustes necesarios | 7 (1 – 1.5 días) |
| Documentación y entrega | Elaboración de manual de uso, documentación técnica del modelo de datos, despliegue en GitHub y presentación final. | 13 (2- 3 días) |

**TABLA 1.**  **Módulos para desarrollar**

El promedio de puntos de esfuerzo utilizados será de 72. Cada módulo incluirá revisiones periódicas, permitiendo retroalimentación tanto del cliente como de la profesora, para al final del semestre entregar un sistema funcional y documentado, con bases de datos que permitan a la empresa de transporte tomar decisiones estratégicas basadas en datos limpios y estructurados. **Luego, como recursos humanos se necesitó de un equipo de desarrollo, conformado por** 4 estudiantes de MACC e Ingeniería Electrónica, encargados del proceso total del proyecto para entregarle a la cliente Sandra Zúñiga. Asimismo, se tuvo un acompañamiento y s**upervisión de la docente Tatiana Cabrera, encargada de la revisión constante del trabajo desarrollado y dar retroalimentaciones claves en búsqueda de mejoras en el trabajo del equipo.** En la siguiente tabla se encuentran las tecnologías y herramientas utilizadas:

|  |  |
| --- | --- |
| Base de datos relacional | MySQL |
| Base de datos no relacional | MongoDB |
| ETL | Excel |
| Backend | Node.js |
| Gestión ágil | Trello |
| Actualizaciones | GitHub |
| Comunicación con cliente | Zoom |

**TABLA 2. Tecnologías del proyecto junto a la herramienta usada por cada una**

**ALTERNATIVAS DE SOLUCIÓN**

**SOLUCIONES EXISTENTES**

**POSIBLES SOLUCIONES POSIBLES**

**ALTERNATIVA DE SOLUCIÓN SELECCIONADA**

Se escogió como solución una aplicación que solo será de uso para administradores y supervisores, esta contará con 6 módulos, siendo estos: cliente, usuario, factura, gastos, conductor, vehículo, viaje. Cada módulo tiene información relevante que se relaciona en factura, donde estará registrada información de viajes como el documento del conductor, la placa del vehículo, el identificador del cliente y lo mas relevante para nuestro cliente las utilidades y los gastos. La idea de esta aplicación es que los usuarios puedan consultar información como multas de un conductor, valor de un Soat, reparaciones de un vehículo etc. Con el fin de tomar decisiones importantes tales como despedir un conductor, vender un vehículo o fijar el precio de determinado viaje basándose en los gastos que este conlleva

Sobre los diferentes módulos: conductores, debe permitir su consulta por documento, accediendo a información como numero de multas, numero de viajes, sin embargo no se tendrá en cuenta la información específica de la fecha de multas; viajes, describe la información del lugar de origen y destino, además de la cantidad de escalas, sin embargo, no especifica los lugares de las escalas; gastos, permite consultar el precio por Id, el precio es promedio y actualizado con la frecuencia que el usuario desee pero al ser un promedio no se adecua a cada zona del país en específico; vehículos, tendrá información sobre el número de reparaciones pero no describirá cuales son. El usuario principalmente consultara el módulo facturas, pues es el que brinda información relacionando vehículos, conductores clientes y gastos, sin embargo, la utilidad total de un viaje se calculara restándole gastos al pago inicial, por lo que se obtendrá un aproximado

**LIMITACIONES**

* No se contempla la localización en tiempo real de vehículos ni pasajeros
* La información de cada viaje es definida por lugar de salida y llegada junto con duración, no se especificará cual es la ruta a tomar
* El cliente de la empresa de transporte no tendrá acceso a la aplicación

**METODOLOGIA**

Se uso la metodología kanban, y le evidencia esta en el siguiente **link**: https://trello.com/invite/b/67a13ac3a9133e5b4ebd7466/ATTIa544d982ce8edae08ba40752628b4cd491A900F2/ing-datos

FOTO DEL TABLERO

ALTERNATIVAS PROPUESTAS

ALTERNATIVA ELEGIDA

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.REQUISITOS FUNCIONALES**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.**

Diagrama

El contenido generado por IA puede ser incorrecto.**DIAGRAMA** **DE CLASES**

**MAPA STAKEHOLDERS**

Imagen que contiene Interfaz de usuario gráfica

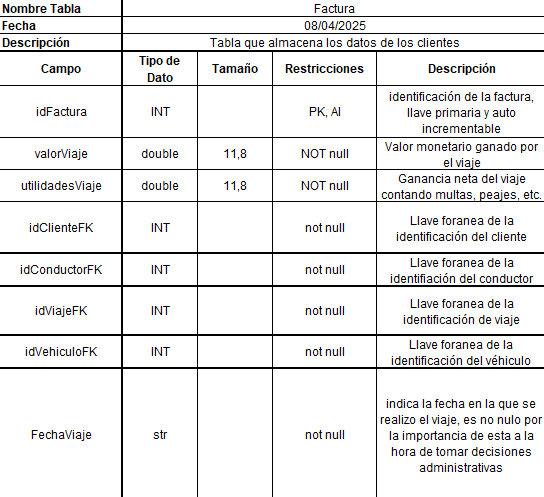
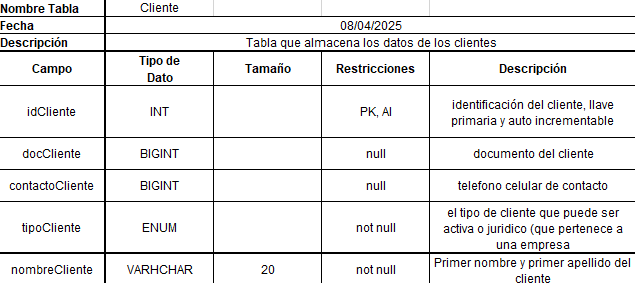
El contenido generado por IA puede ser incorrecto.

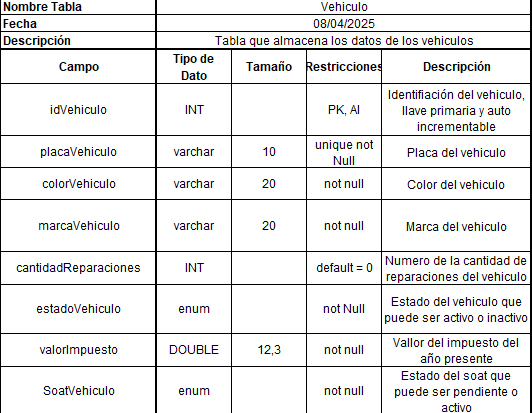
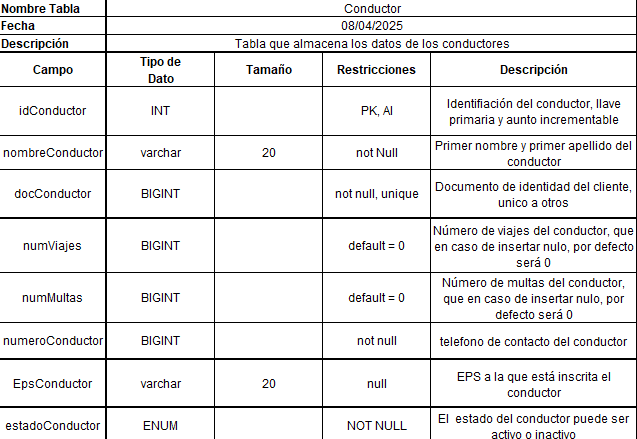
EXPLICACION

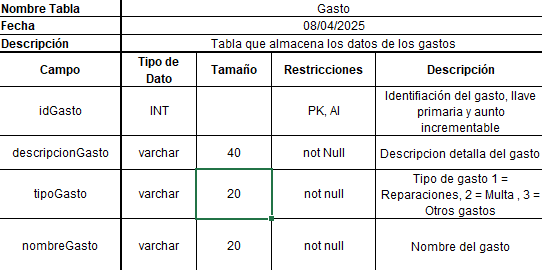
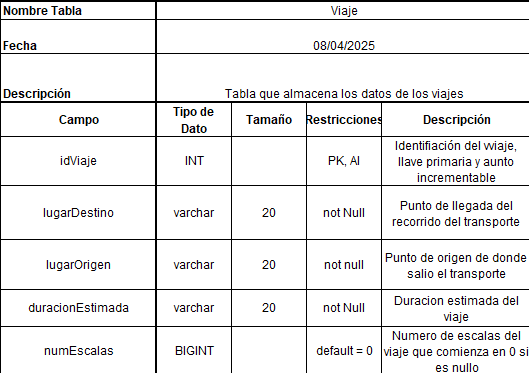
**MODELO CONCEPTUAL**

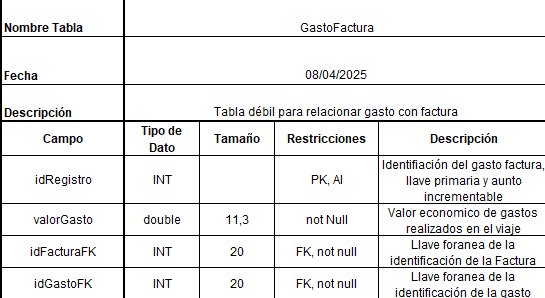
**ACTUALIZAR EL QUE TENEMOS**

**DICCIONARIO DE DATOS**









**IMPLEMENTACION DE LA SOLUCION EN MYSQL**

Para implementar la solución creamos las tablas descritas en el diagrama de clases, una vez se realizó la inserción de 50 datos en cada una de ellas se verifico que este proceso se hubiera realizado de manera correcta realizando las consultas básicas a todas las tablas, el resultado de esto fue el siguiente:

Tabla conductor:

Tabla

El contenido generado por IA puede ser incorrecto.

Tabla Vehículo:

Tabla

El contenido generado por IA puede ser incorrecto.

Tabla Cliente

Tabla

El contenido generado por IA puede ser incorrecto.

Tabla Viaje

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Tabla Gasto

Texto, Tabla

El contenido generado por IA puede ser incorrecto.

Tabla factura

Tabla

El contenido generado por IA puede ser incorrecto.

Tabla GastoFactura

Tabla

El contenido generado por IA puede ser incorrecto.

Sobre las consultas especificas a continuación se mostraran las 7 más relevantes que son: consultar clientes que sean de tipo jurídico, consultar conductores con alguna multa, consultar el viaje con mayor número de escalas, consultar gastos correspondientes a multas, consultar gastos correspondientes a reparaciones del vehículo, consultar conductores inactivos y consultar los 5 conductores con mayor número de viajes.

script consulta 1:

Texto

El contenido generado por IA puede ser incorrecto.

Resultado consulta 1

Interfaz de usuario gráfica, Aplicación, Tabla

El contenido generado por IA puede ser incorrecto.

script consulta 2:



Resultado consulta 2:

Interfaz de usuario gráfica, Tabla

El contenido generado por IA puede ser incorrecto.

script consulta 3:



Resultado Consulta 3:



Script consulta 4:

Texto

El contenido generado por IA puede ser incorrecto.

Resultado consulta 4:

Texto

El contenido generado por IA puede ser incorrecto.

Script 5:

Texto

El contenido generado por IA puede ser incorrecto.

Resultado consulta 5:

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

Script 6:

Imagen de la pantalla de un celular con letras

El contenido generado por IA puede ser incorrecto.

Resultado consulta 6:

Tabla

El contenido generado por IA puede ser incorrecto.

Script 7:

Texto

El contenido generado por IA puede ser incorrecto.

Resultado Script 7:

Interfaz de usuario gráfica, Tabla

El contenido generado por IA puede ser incorrecto.

MODIFICACIONES

En el caso de las modificaciones la primera que surgió fue la de agregar el campo Eps a la tabla conductores puesto que fue un campo sugerido por la docente que nos pareció bastante correcto pero que no estaba en la tabla creada inicialmente

Script:



Ocurrió algo similar con la tabla gasto, donde estaba registrado un nombre conciso del gasto para poder identificarlo de manera sencilla pero en caso de necesitar una mayor profundidad, como por ejemplo el nombre “frenos” era fácil de identificar como un gasto referente a vehículo pero difícil saber si era un cambio total o un ajuste, para esto se agregó el campo “descripción gasto”



Las demás modificaciones que surgieron se realizaron porque había registros realizados de manera incorrecta y era mas eficiente realizar un update a volver a realizar las inserciones desde 0.

Script 3:

Imagen que contiene Escala de tiempo

El contenido generado por IA puede ser incorrecto.

Script 4:

Texto

El contenido generado por IA puede ser incorrecto.

Script 5:

Texto

El contenido generado por IA puede ser incorrecto.

Además, se realizó la eliminación de 2 registros que estaban duplicados:

Texto

El contenido generado por IA puede ser incorrecto.

Para las consultas multitablas la mayoría de estas se realizaron en la tabla factura, ya que como se menciono anteriormente esta es la que relaciona la mayor cantidad de clientes, sin embargo, otro parte se realizo en la tabla GastoFactura, Tabla que surgió durante la normalización y contenía información específica sobre el valor de los gastos que surgian en determinado viaje

Script 1:

Texto

El contenido generado por IA puede ser incorrecto.

Script 2:

Texto

El contenido generado por IA puede ser incorrecto.

Resultado script 2:



Además de estas consultas se realizaron procedimientos almacenados, vistas y triggers que agilizaban consultas, actualizaciones o modificaciones automáticas que serían parte del día a día de nuestro cliente.

Script 1, procedimiento para consultar cual es el gasto que mayor peso tuvo en la factura.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Script 2, procedimiento para consultar toda la informacion de un conductor especifico.

Texto

El contenido generado por IA puede ser incorrecto.

Script 3,procedimiento para actualizar la eps de determinado conductor.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Script 4, procedimiento para actualizar el estado de un conductor. Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Script 5, procedimiento para cambiar el estado del SOAT de determinado vehiculo

Texto

El contenido generado por IA puede ser incorrecto.

Script 6, vista que desde factura permite consultar la informacion completa, con nombre del conductor los clientes, placa del vehiculo etc.

Texto

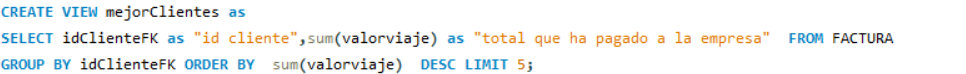
El contenido generado por IA puede ser incorrecto.

Resultado script 6:

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Script 7, vista que permite consultar cuales son los clientes que han aportado más dinero a la empresa

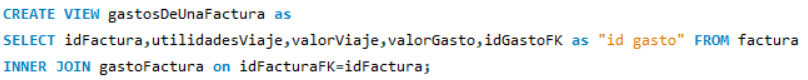


Resultado script 7:

Tabla

El contenido generado por IA puede ser incorrecto.

Script 8, vista que permite consultar todos los gastos de una factura incluyendo su valor, descripcion y tipo



Resultado script 8:

Tabla

El contenido generado por IA puede ser incorrecto.

Script 9, trigger que de manera automática añade 1 al numero de reparaciones a un vehículo una vez se realice una inserción con tipo de gasto 1 en gastoFactura

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

Script 9, trigger que de manera automática añade 1 al numero de multas de un conductor una vez se realice una inserción con tipo de gasto 2 en gastoFactura

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

**REFERENCIAS**