**PROYECTO INGENIERIA DE DATOS**

**EMPRESA INVERTRANS**

**INTEGRANTES:**

Héctor Rafael León Molina

Juan Nicolas Arévalo Lemus

William Alejandro Ramírez Montaña

Juan Sebastian Arroyo Zúñiga

**Bogotá D.C**

**25/02/2025**

**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.*

*Es por eso 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 mas manejable. Esto combinado con un workshop -un espacio de dialogo 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**

Nuestro cliente posee una base de datos en Excel, que, a pesar de tener información completa, no permite consultas eficientes, ya que tiene datos duplicados lo cual complica tomar decisiones basadas en la información obtenida por la poca fiabilidad de esta.

**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 numero 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 a 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 SIMILARES**

**SOLUCIONES POSIBLES**

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.**

**aaa**

**DIAGRAMA**  **DE CLASES**

Diagrama

El contenido generado por IA puede ser incorrecto.

**MAPA STAKEHOLDERS**

Imagen que contiene Interfaz de usuario gráfica

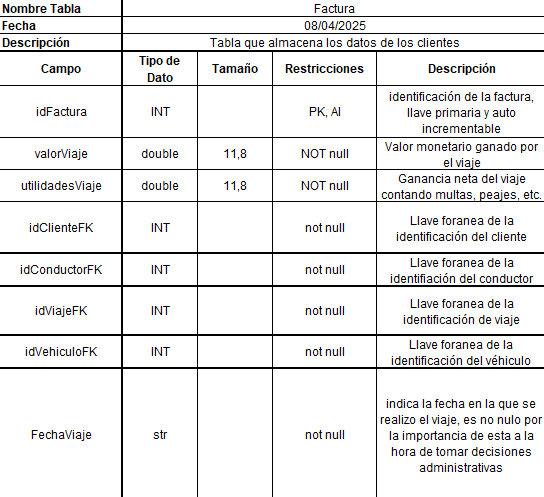
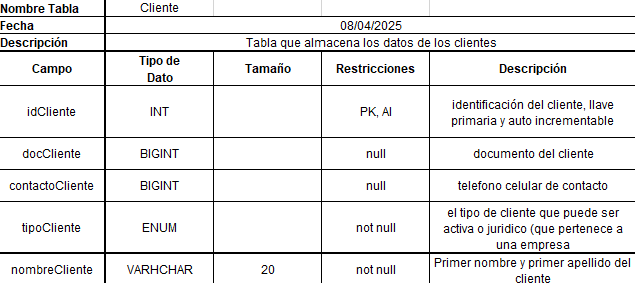
El contenido generado por IA puede ser incorrecto.

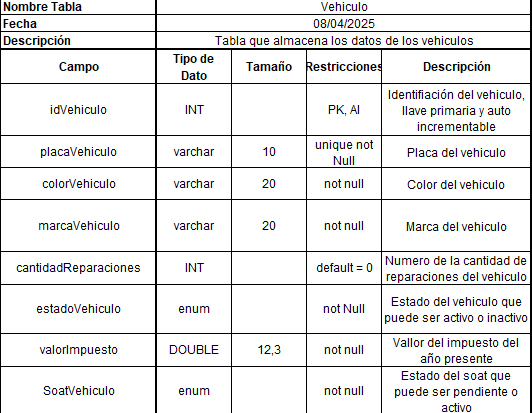
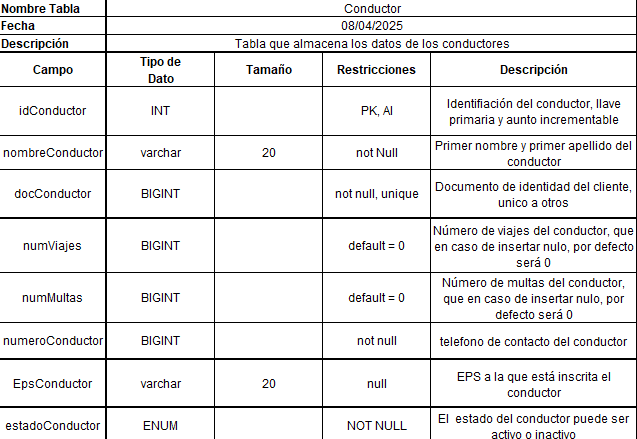
EXPLICACION

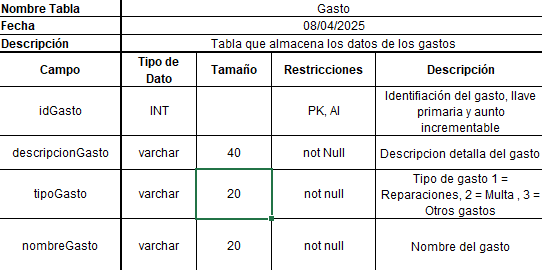
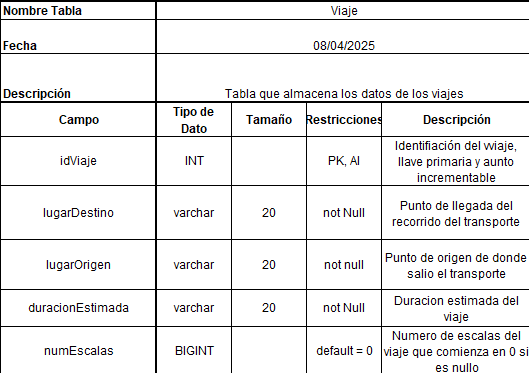
**MODELO CONCEPTUAL**

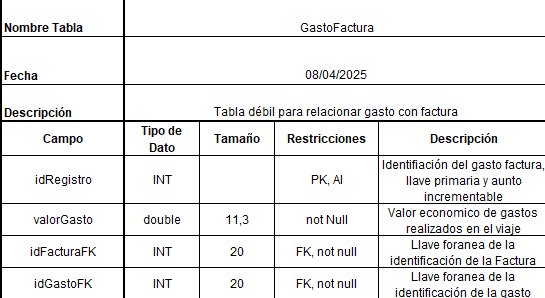
**ACTUALIZAR EL QUE TENEMOS**

**DICCIONARIO DE DATOS**









**REFERENCIAS**