UNIVERSIDAD NACIONAL DE COLOMBIA

CURSO DE BASES DE DATOS

2016353 – 2

EJERCICIO NÚMERO 3

Grupo de trabajo número 21

Index, primary key y agrupación de datos

Por:

Bonilla Muñoz Daniel Mauricio

Garzon Contreras Nestor Sebastian

Rojas Garzón María Alejandra

Silva Monroy David

Bogotá, 08 de octubre 2020

TABLA DE CONTENIDO

[**OBJETIVO DEL EJERCICIO 1**](#_heading=h.gjdgxs)

[**ENUNCIADO DEL PROBLEMA 1**](#_heading=h.30j0zll)

[**MARCO TEÓRICO 1**](#_heading=h.1fob9te)

[**METODOLOGÍA PARA LA SOLUCIÓN DEL PROBLEMA 1**](#_heading=h.3znysh7)

[**DIAGRAMAS DE FLUJO PARA LA SOLUCIÓN DEL PROBLEMA 1**](#_heading=h.tyjcwt)

[**DIAGRAMAS DE FLUJO ESTRUCTURADOS PARA LA IMPLEMENTACIÓN 1**](#_heading=h.3dy6vkm)

[**MODELO CONCEPTUAL 1**](#_heading=h.1t3h5sf)

[**MODELO LÓGICO 1**](#_heading=h.4d34og8)

[**MODELO FÍSICO 1**](#_heading=h.2s8eyo1)

[**MUESTRA DE RESULTADOS 2**](#_heading=h.17dp8vu)

[**Resultados utilizando la base de datos ASE SAP 2**](#_heading=h.3rdcrjn)

[**Resultados utilizando la base de datos ORACLE 2**](#_heading=h.26in1rg)

[**Resultados utilizando la base de datos MS SQL Server 2**](#_heading=h.lnxbz9)

[**CONCLUSIONES 2**](#_heading=h.35nkun2)

[**RECOMENDACIONES 2**](#_heading=h.1ksv4uv)

[**BIBLIOGRAFÍA 2**](#_heading=h.44sinio)

# OBJETIVO DEL EJERCICIO

Desarrollar una base de datos relacional para la empresa Ventas Colombia SAS poniendo en práctica los conceptos vistos en el curso hasta el momento identificando de los requerimientos propuestos cuales se pueden llevar a cabo. Todo esto desarrollado en tres motores de bases de datos: Oracle, SAP y Microsoft SQL Server

# ENUNCIADO DEL PROBLEMA

¿Cómo desarrollar una base de datos relacional para la empresa Ventas Colombia SAS que es una empresa dedicada a vender al mayor productos que adquiere por contratos con los proveedores dados los requerimientos en múltiples motores de bases de datos para que sea un sistema confiable, los datos sean consistentes y estén disponible de manera segura y controlada?

# MARCO TEÓRICO

**DATO:** Es un componente que describe el manejo de la información y representación de la misma, esta información es usada posteriormente en la ejecución de un algoritmo. Tipos de datos: caracteres, numéricos, imágenes, fechas, monedas, texto, bit, decimales y varchar.

**Tablas:** Las tablas son objetos que contienen todos los datos estos se organizan por medio de filas y columnas similar a una hoja de cálculo. Las filas representan un registro único y las columnas representan un campo específico.

Además hay varios tipos de tablas , estas características que las diferencian entre ellas dependen del servidor en el cual estemos trabajando. Cada clase de tabla está diseñada para llevar a cabo diferentes tareas y cuentan con funciones únicas.

Los principales tipos de tablas son:

* Tablas con particiones: Estas tablas se dividen de manera horizontal sus componentes, creando distintas particiones con la capacidad de poder guardar archivos de manera segmentada, y agrupándolos en diferentes lugares de la base de datos.

Al tener una tabla particionada obtenemos una mayor organización, más control en la información , y nos da la posibilidad de acceder de manera directa a un punto en específico, aumentando la eficacia y la velocidad de trabajo en la base de datos.

En las tablas con particiones la integridad global entre los archivos, todavía se mantiene y toda la información se encuentra en el mismo lugar.

* Tablas temporales: Las tablas temporales destacan por el tiempo de disponibilidad, estas se dividen en 2 clases donde cada una posee diferencias de visibilidad y tiempo de funcionamiento.

Primero- Tablas locales, solo podrán ser vistas por el usuario que se encuentre conectado en ese momento, y desaparece cuando este se desconecta del servidor de bases de datos.

Segundo-Tablas globales, será visible en el momento en que todos los usuarios estén conectados desde su creación hasta su eliminación cuando todos los usuarios se desconecten.

**ASE SAP:** Es un motor de bases de datos de la empresa Sybase, antes llamado Sybase SQL server, tuvo que cambiar de nombre en 1996 para evitar la confusión con su competencia Microsoft SQL server.

**ORACLE DATABASE:** Es una de las principales bases de datos de la compañía Oracle Corporation junto con MySQL database y Oracle NoSQL database.

**MICROSOFT SQL SERVER:** Es un gestor de base de datos desarrollado por la empresa Microsoft, el cual permite trabajar de modo cliente servidor, donde la información se almacena en el servidor y los usuarios y clientes tienen que acceder a ella.

**REGLAS DE CODD:** Creadas por Edgar Frank Codd, con el fin de brindar una serie de parámetros que una base de datos debe cumplir para poder considerarse una base de datos relacional. Las trece reglas son:

Regla número cero (Regla de fundación): Toda la información de la base de datos, ya sea usuarios, tablas, índices deben gestionarse mediante sus capacidades relacionales dentro de la misma base de datos.

Regla número uno (Regla de la información): Toda la información de la base de datos se representa y almacena en tablas.

Regla número dos (Regla de acceso garantizado): Todos los datos deben ser claros sin presentar ambigüedad, además tienen que ser lógicamente direccionables de tal modo que se especifique la tabla que lo contiene, la columna y la llave primaria.

Regla número tres (Regla de tratamiento de valores nulos): El sistema de gestión de la base de datos debe tener una representación para la información inaplicable y faltante, para ello se establecen campos nulos.

Regla número cuatro (Regla de Catálogo):Los usuarios autorizados, por medio de un catálogo en línea basado en el modelo relacional, deben poder acceder a la estructura de la base de datos.

Regla número cinco (Regla del sublenguaje de los datos): El sistema debe soportar un lenguaje relacional con las siguientes características:

-Con sintaxis lineal.

-Se pueda utilizar de forma interactiva.

-Soporte operaciones tales como:

* Definición de datos.
* Manipulación de datos.
* Actualización y recuperación.
* Seguridad e integridad.
* operaciones de administración de transacciones.

Regla número seis (Regla de actualización): Todas las vistas deben ser actualizadas por el sistema.

Regla número siete (Regla de inserción, actualización y borrado): El sistema debe poder realizar manipulación de datos sobre conjuntos de registros, esto significa por actualizar más de un registro a la vez.

Regla número ocho (Regla de independencia física de los datos): Si se realiza un cambio en los métodos de acceso y representación en un cliente, este queda inalterado a nivel lógico.

Regla número nueve (Regla de independencia lógica de los datos): Si se realizan cambios en las tablas que almacenan la información la lógica de los datos permanece inalterable.

Regla número diez (Regla de independencia de la integridad):Las reglas de integridad tienen que estar almacenadas en la base de datos y especificadas por separado, además se deben poder cambiar sin alterar las aplicaciones existentes.

Regla número once (Regla de independencia de distribución): Las distribuciones de las partes de la base de datos tienen que ser invisibles para los usuarios. Estos usos deben seguir funcionando con éxito:

* Cuando se introduce por primera vez una versión distribuida de DBMS.
* Cuando los datos existentes se distribuyen en el sistema.

Regla número doce (Regla de la no subversión): Si el sistema utiliza un lenguaje de bajo nivel para la interfaz además del lenguaje relacional SQL, se debe garantizar que ese lenguaje de bajo nivel no esté en la capacidad de subvertir el sistema.

**Campos:** Es un espacio de almacenamiento para un dato en particular .Los campos definen los tipos de datos que limitarán lo que se podrá almacenarse en él, además es la mínima información a la que se puede acceder.

**Registros:** Los registros son conjuntos de campos que contienen datos que pertenecen a una misma entidad, se le asigna automáticamente un número consecutivo que en varias ocasiones es usado como índice .

**Vistas:** Es una tabla virtual donde su contenido es definido por una consultas, y actúa como un filtro de las tablas subyacentes a las que se hace referencia en ella. Las vistas suelen usarse para centrar, simplificar y personalizar la percepción de la base de datos para cada usuario además se puede usar como una mecanismo de seguridad, los usuarios que usen vistas podrán acceder a los datos de sus tablas pero tendrán restricciones a la hora de manipular datos.

**Secuencias:** se usa para generar automáticamente valores enteros secuenciales únicos y asignarlos a campos numéricos. Podemos asignarle un mínimo donde inicia la secuencia y un máximo que no sobrepase, además de configurar si queremos que sea cíclico cuando llegue al valor máximo.

Las secuencias son independientes de las tablas, pero se utilizan generalmente para una tabla específica .

Las sentencias “nextval” y “currval” nos ayudarán a recuperar los valores de la secuencia creada, La primera vez que usemos “nextval”, se inicializa la secuencia en el valor que le indiquemos, luego incrementa la secuencia y nos retorna el valor. Para recuperar el valor actual se usa la sentencia “currval”. Usar secuencias ayuda a reducir la cantidad de código a escribir.

**Atributo identity:** Los campos numéricos pueden tener atributos de tipo identity.Los valores de este campo generan valores secuenciales que se inician en 1 y se incrementa en 1 automáticamente. Este tipo de datos se llena automáticamente y no es necesario usar otra sentencia.

**Table scan:** Este tipo de escaneo consiste en leer la tabla entera ( filas y columnas) tal y como se almacena en el disco. siendo una de las operaciones más costosas.

**Índices:** Es un objeto definido sobre una columna de tabla, que permite localizar de forma rápida las filas de la tabla en base a su contenido en la columna indexada. Mejorando la velocidad de las operaciones, lo hace por medio de un identificador único de cada fila de la tabla, permitiendo un rápido acceso a los registros. Su funcionamiento es muy parecido al índice de un libro.

Su creación requiere su propio espacio en disco y contiene una copia de los datos de la tabla.

* **Tipos de índices:**

**Índice simple y compuesto:** Un índice simple está definido sobre una sola columna de la tabla. Un índice compuesto está definido por varias columnas de la misma tabla. Cuando se crea un índice los registros aparecerán ordenados por el campo indexado.

**Índice agrupado y no agrupado:** Un índice agrupado define el orden en el cual los datos son físicamente almacenados en una tabla. En una tabla solo se puede ordenar de una sola forma, por esta razón no puede haber más de un índice agrupado por tabla. Cuando creamos un campo con clave primaria en sql server, este crea un índice agrupado automáticamente en esa columna.

También se pueden crear índices agrupados personalizados, para ello tienes que usar la palabra reservada “CLUSTERED” antes de “INDEX”. Si una tabla no tiene un índice agrupado, entonces sus fila de datos estarán almacenadas en una estructura sin ordenar denominada montón.

Los índices no agrupados no ordena los datos de una tabla físicamente. están almacenados en un lugar diferente al de la tabla es similar cuando tenemos un libro y este tiene el contenido en un lugar y el índice en otro. permitiendo tener más de un índice por tabla.

# METODOLOGÍA PARA LA SOLUCIÓN DEL PROBLEMA

Requerimientos:

* Tener un ordenador con conexión estable a internet
* Tener configuradas los 3 motores de bases de datos
* Disponer de los requerimientos para el desarrollo de la bases de datos

Paso 1: Leer los requerimientos dados

Paso 2: Seleccionar requerimientos realizables para la entrega presente

Paso 3: Abstraer de los requerimientos las entidades de las cuales se van a desarrollar las tablas

Paso 4: Identificar qué tipo de dato van a tener cada atributo de las entidades

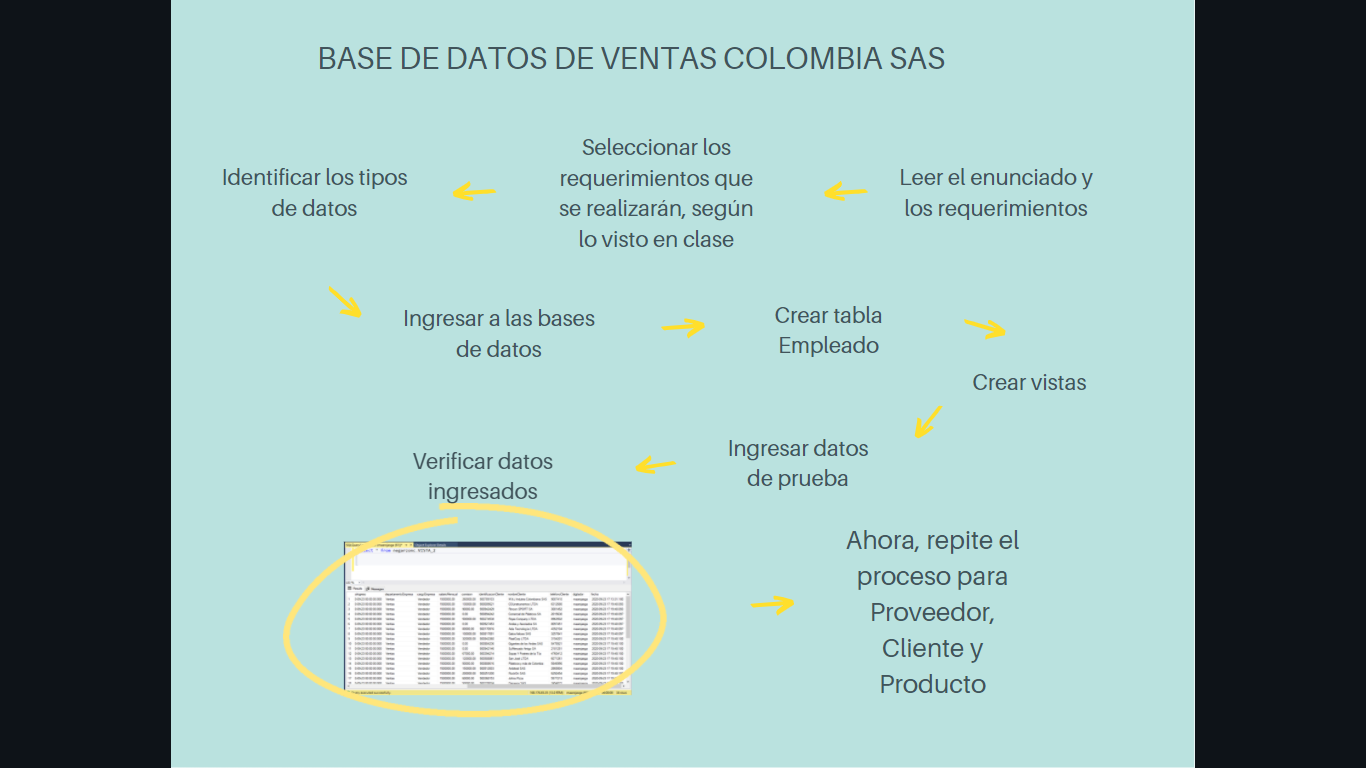
Paso 5: Realizar las diferentes tablas en los motores de base de datos

Paso 6: Crear dos vistas por cada tabla

Paso 7: Otorgar acceso a los otros miembros del grupo en las vistas para que accedan e inserten datos en las tablas.

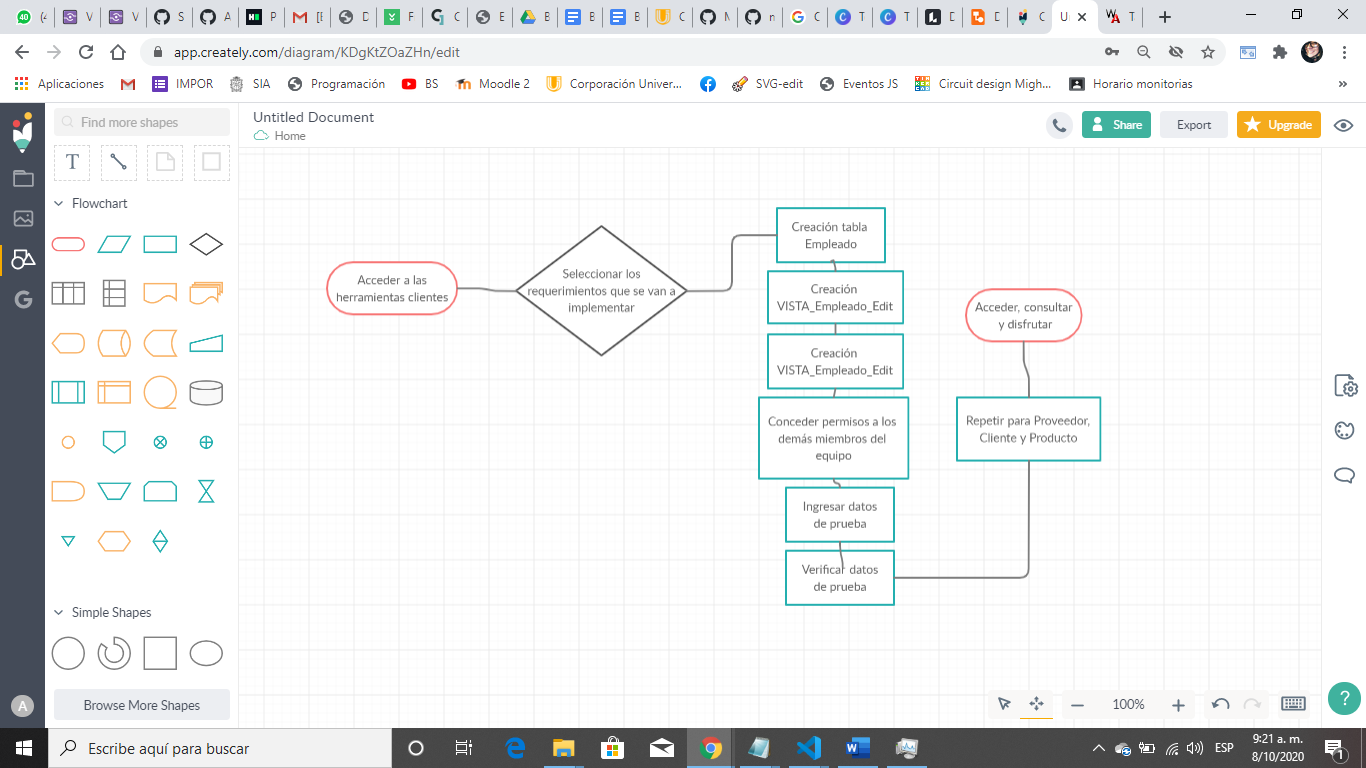
Paso 8: Insertar datos en las tablas.

# DIAGRAMAS DE FLUJO PARA LA SOLUCIÓN DEL PROBLEMA



**Figura 1.** Flujo solución problema creación base de datos

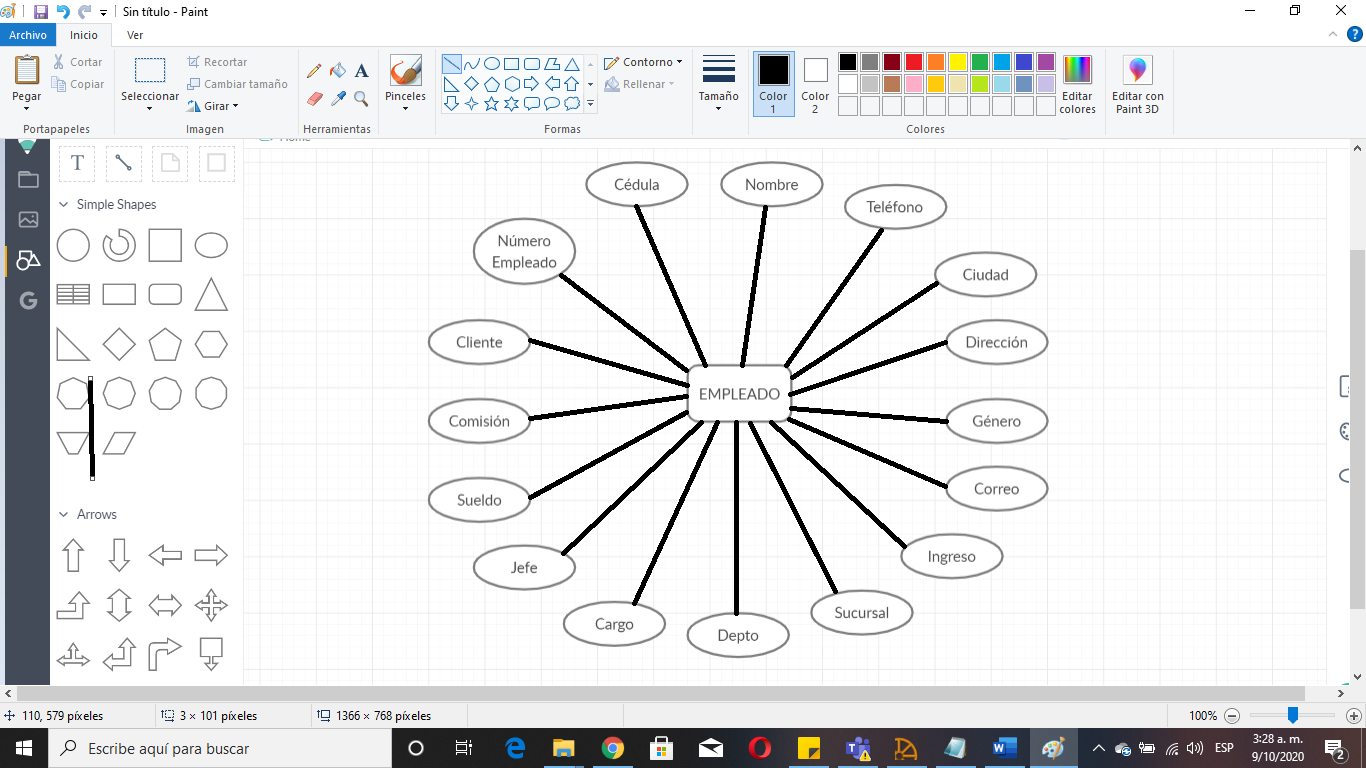
# DIAGRAMAS DE FLUJO ESTRUCTURADOS PARA LA IMPLEMENTACIÓN



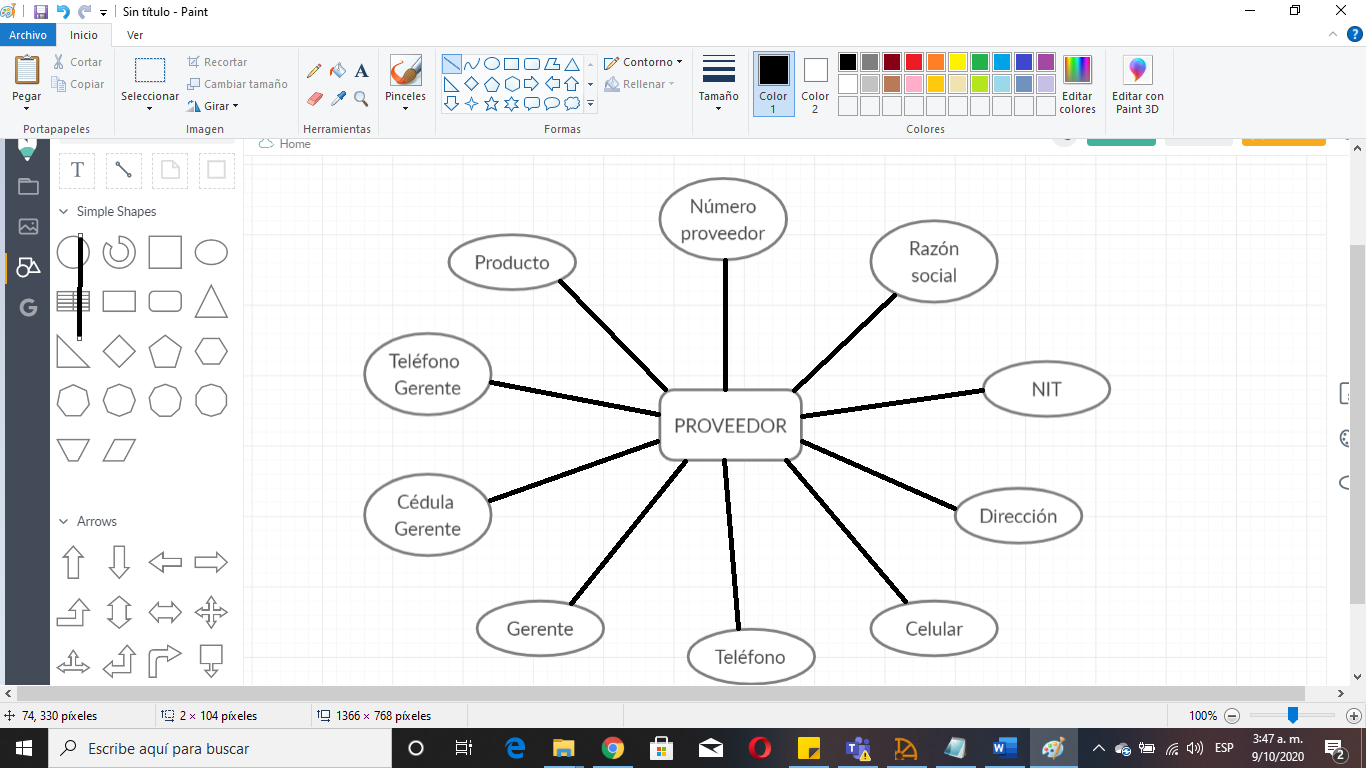
**Figura 2.** Diagrama de flujo estructurado creación de bases de datos.

# 

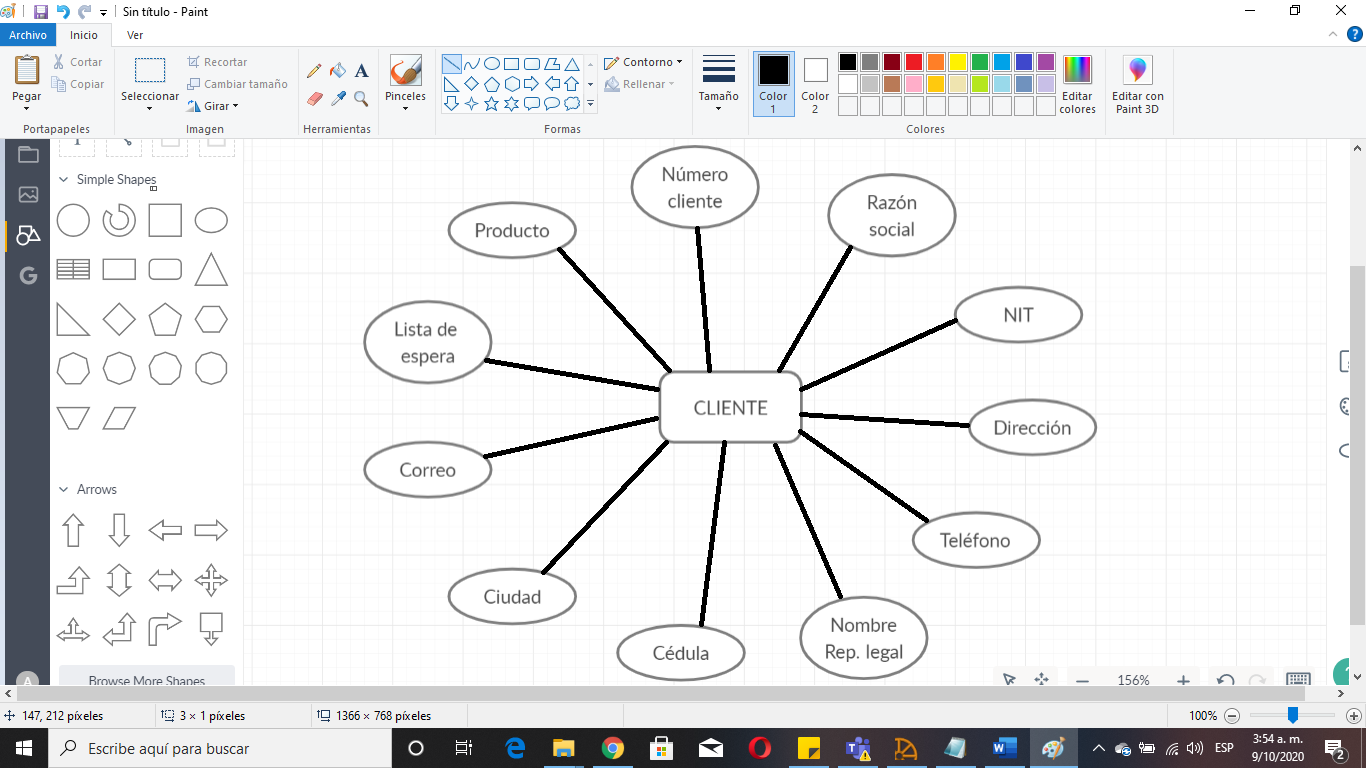
# MODELO CONCEPTUAL



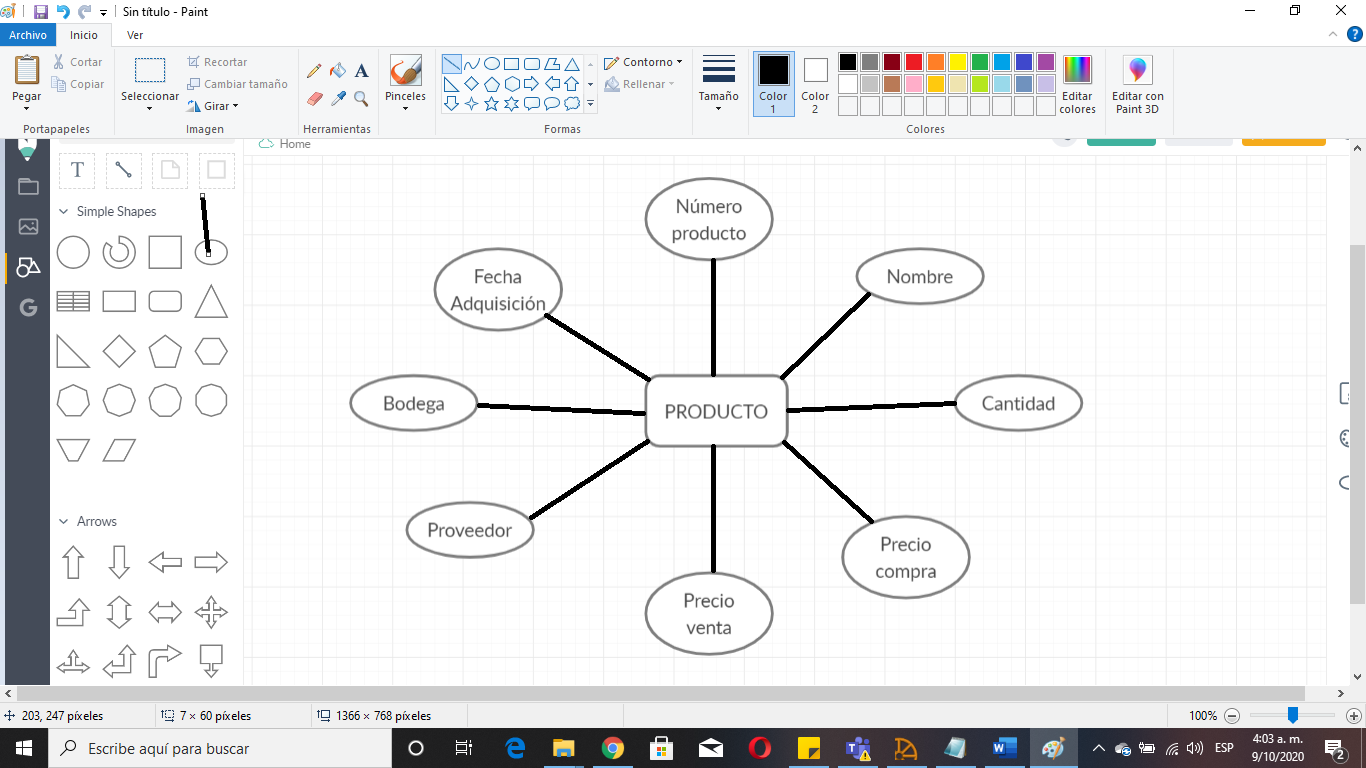
**Figura 3.** Modelo conceptual Empleado



**Figura 4.** Modelo conceptual Proveedor

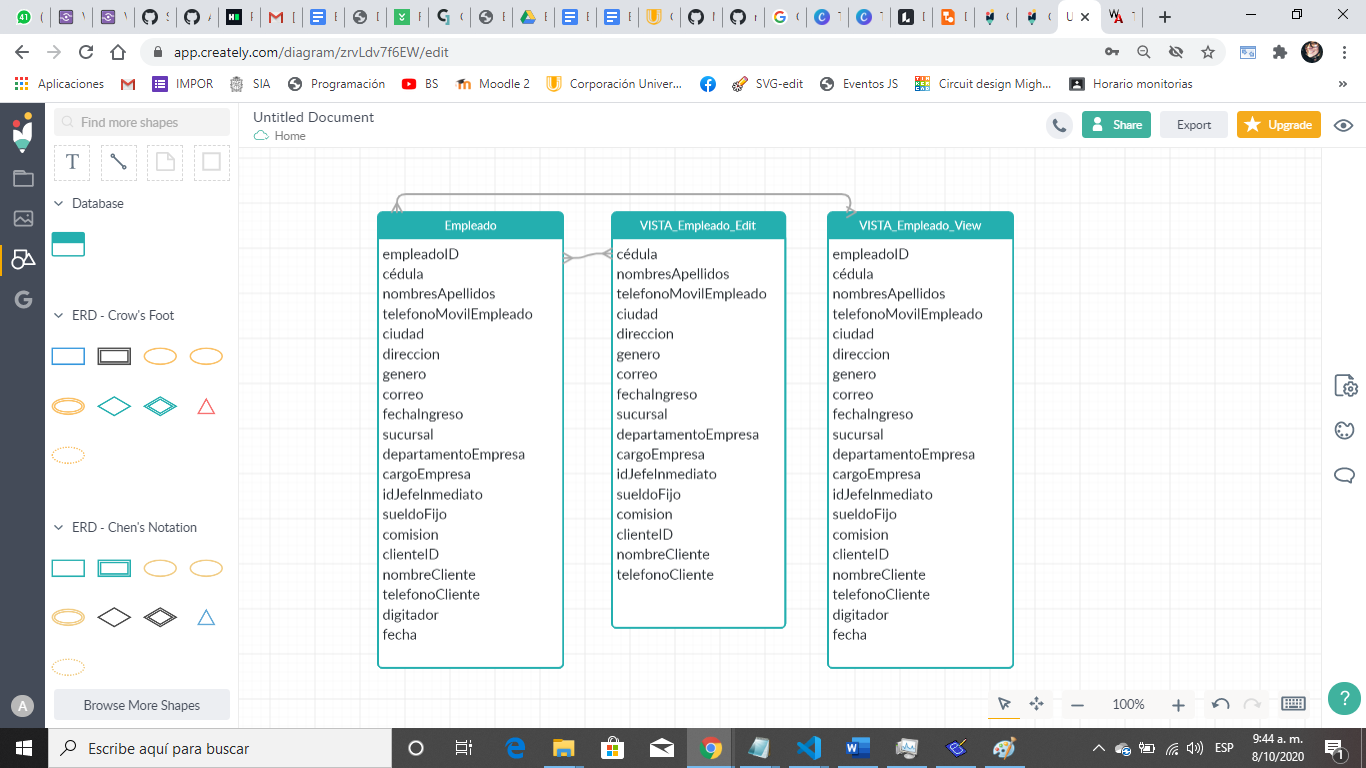


**Figura 5.** Modelo conceptual Cliente

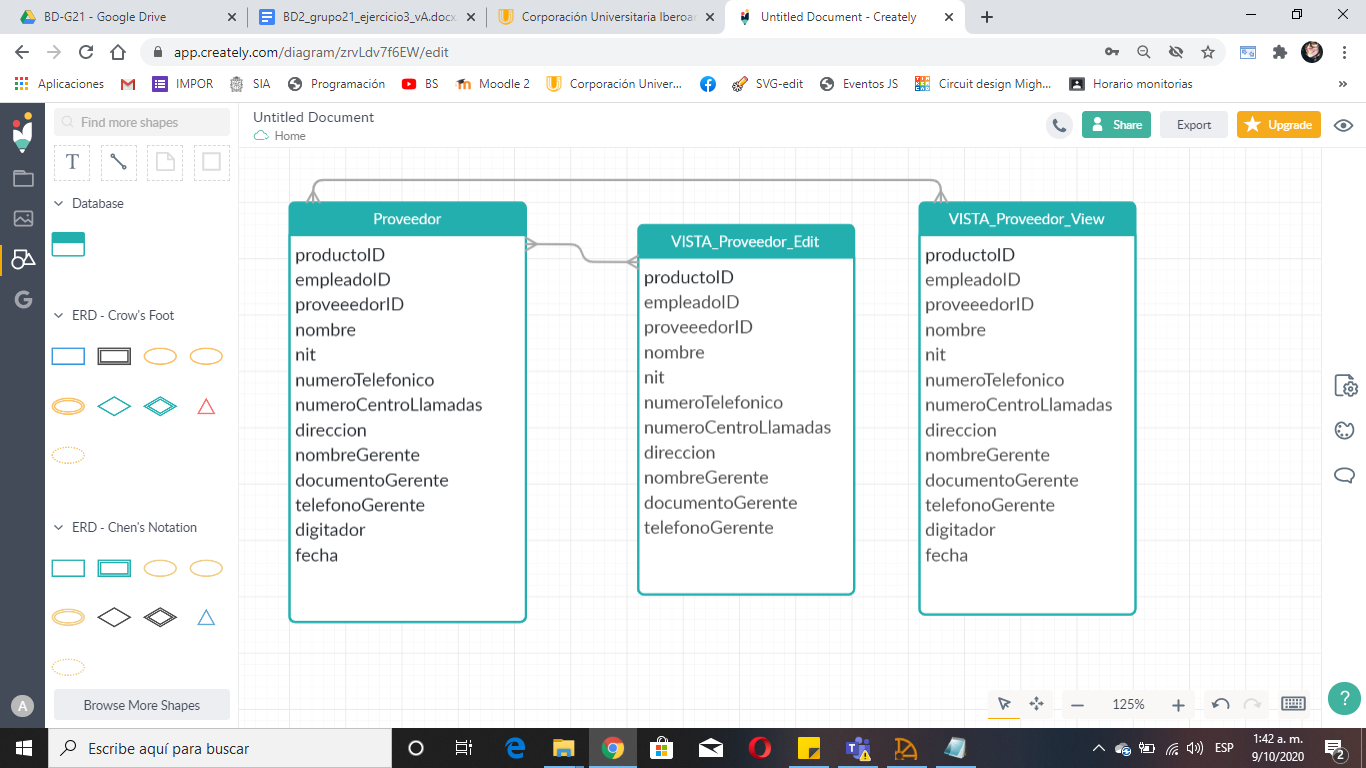


**Figura 6.** Modelo conceptual Producto

# MODELO LÓGICO



**Figura 7.** Modelo lógico tabla empleados



**Figura 8.** Modelo lógico tabla Proveedores

# 

**Figura 9.** Modelo lógico tabla Cliente

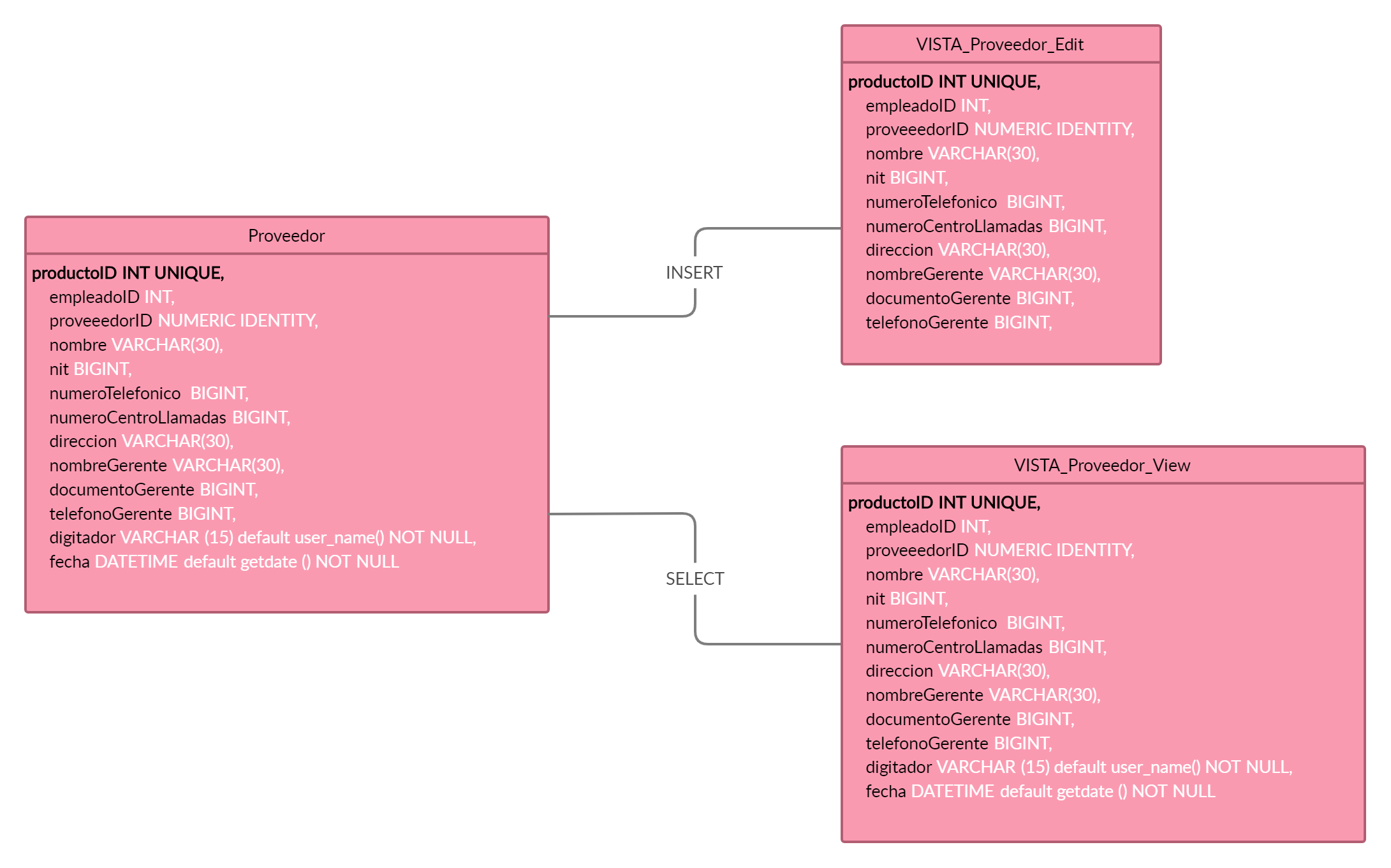


**Figura 10.** Modelo lógico tabla Producto

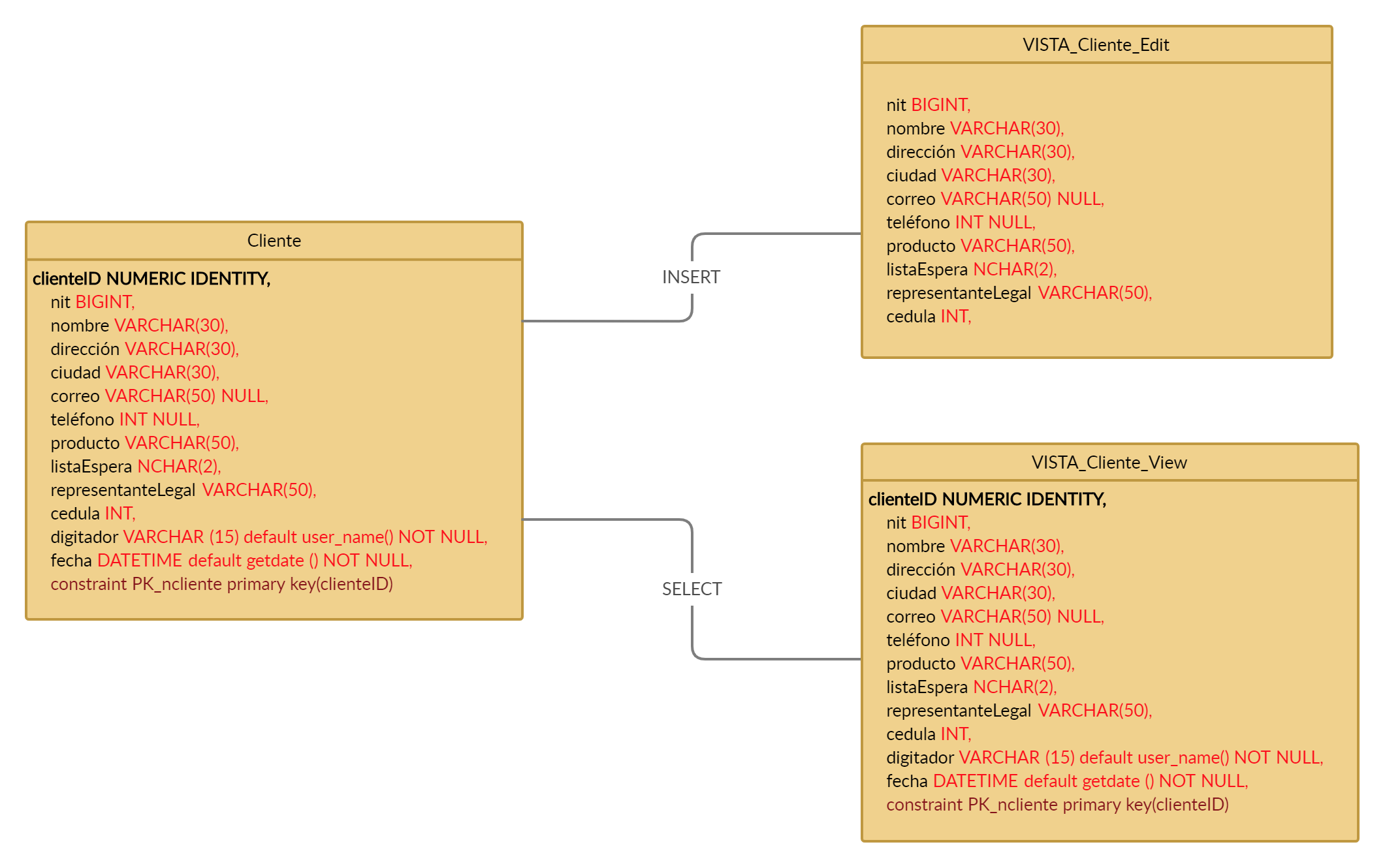
# MODELO FÍSICO



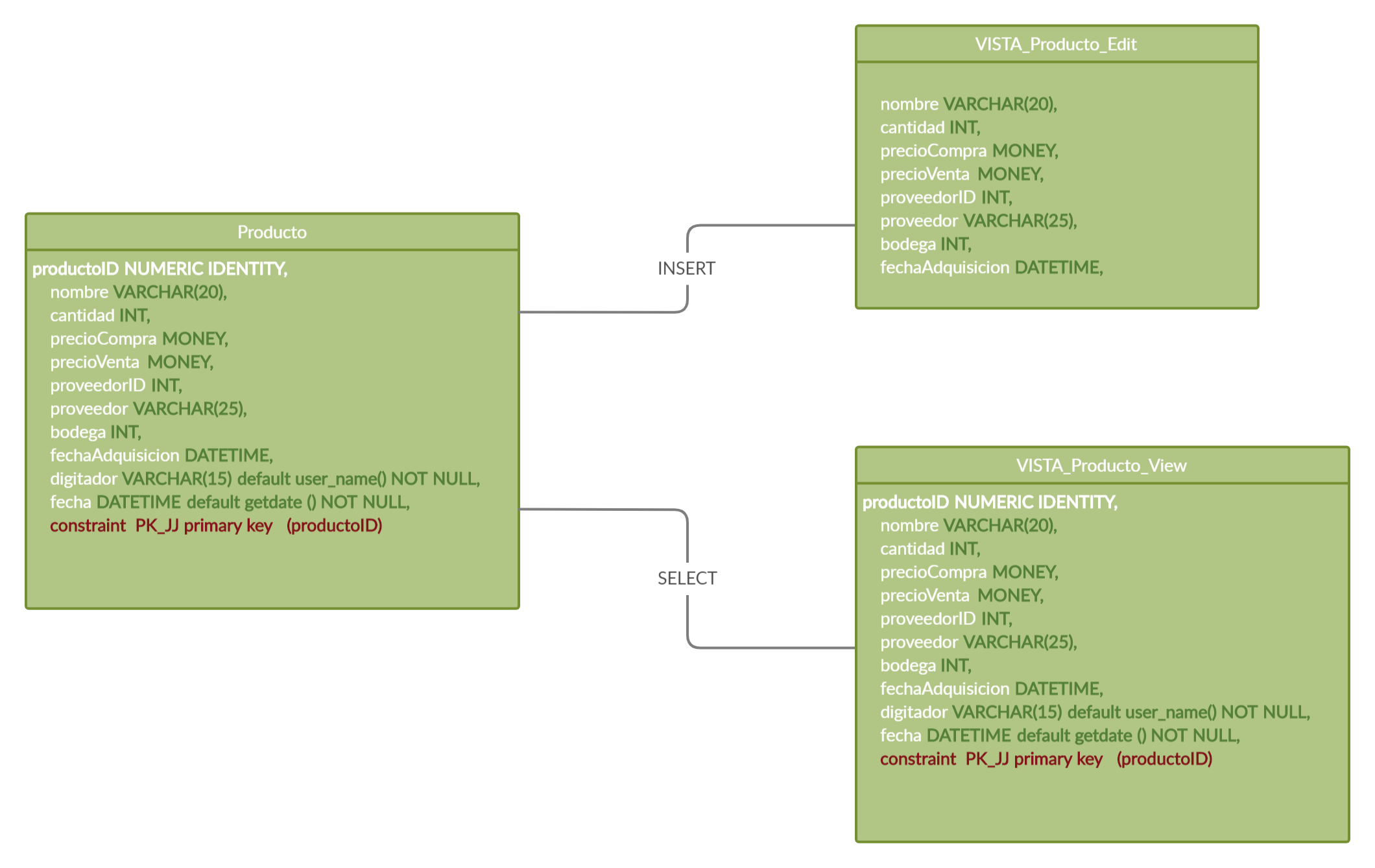
**Figura 11.** Modelo físico tabla Empleado



**Figura 12.** Modelo físico tabla Proveedor



**Figura 13.** Modelo físico tabla Cliente



**Figura 14.** Modelo físico tabla Producto

# MUESTRA DE RESULTADOS

## Resultados utilizando la base de datos ASE SAP



**Figura 15.** Captura de pantalla tablas creadas, índices y las columnas de la tabla Empleado

## Resultados utilizando la base de datos ORACLE

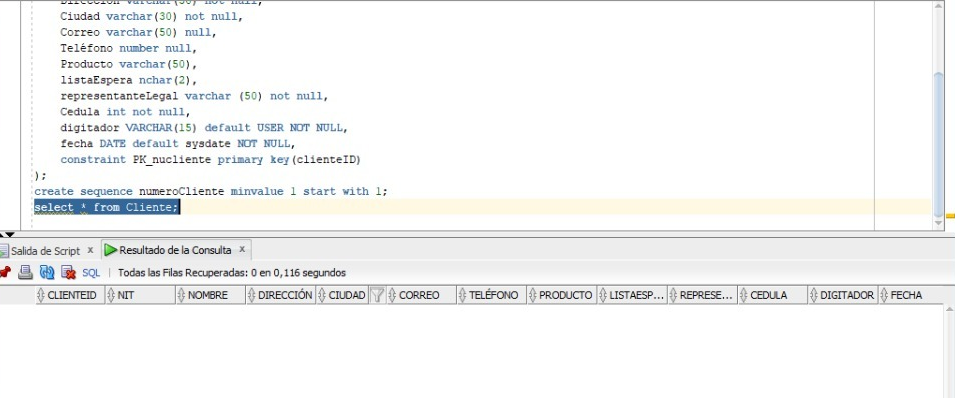


Figura 9. Captura de pantalla al ejecutar sentencias

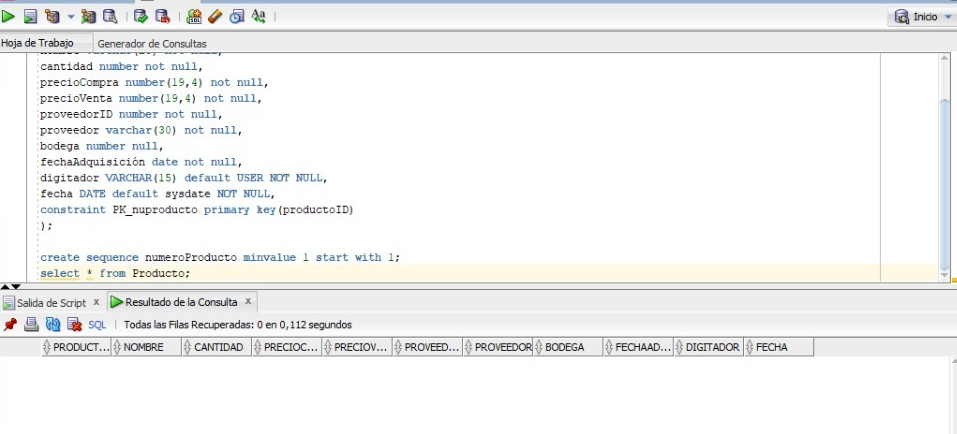


Figura 10. Captura de pantalla al ejecutar sentencias

## Resultados utilizando la base de datos MS SQL Server

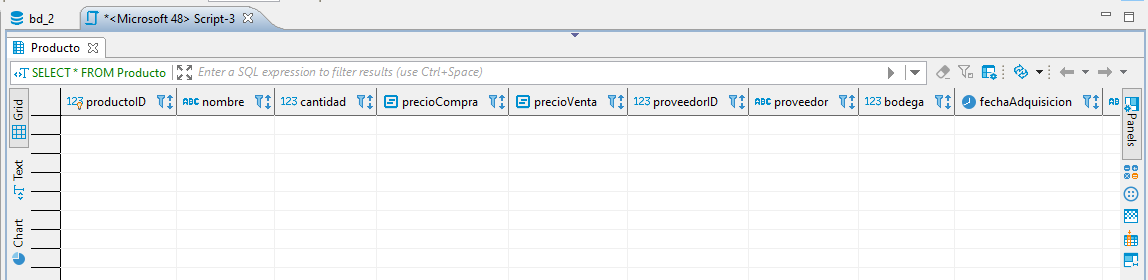


Figura 11. Captura de pantalla al ejecutar sentencias tabla producto

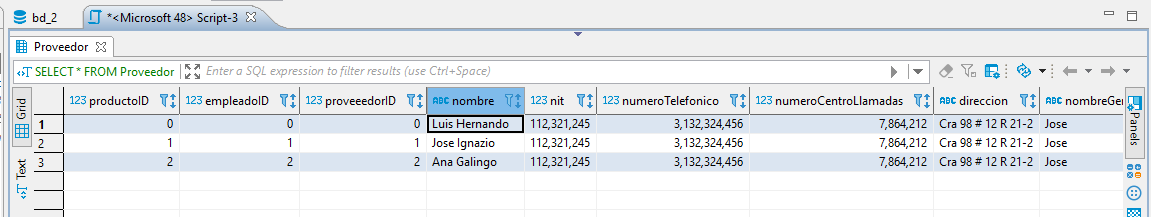


Figura 12. Captura de pantalla al ejecutar sentencias tabla proveedor

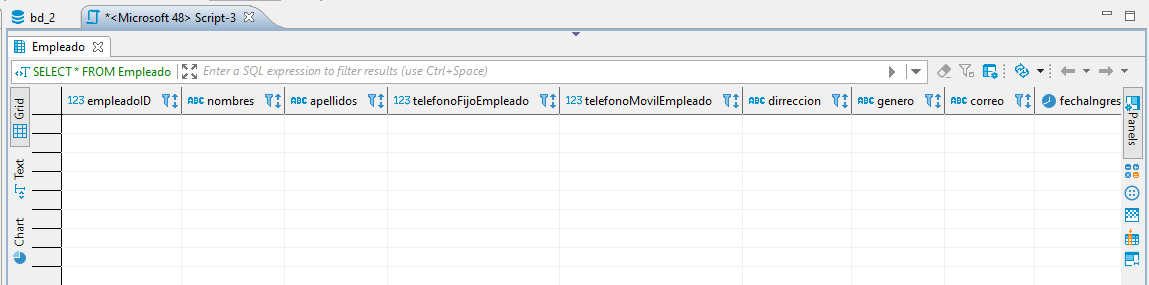


Figura 13. Captura de pantalla al ejecutar sentencias tabla empleado

# CONCLUSIONES

Es importante resaltar la importancia de implementar un base de datos relacional para este ejercicio, ya que facilita el manejo de datos para la empresa “Ventas Colombia SAS” teniendo en cuenta la lista de requerimientos básicos y técnicos dada al equipo de trabajo.

Si bien no disponemos del conocimiento para cumplir todos los requerimientos necesarios para la realización de la base de datos de la empresa, y a pesar del escaso tiempo para el desarrollo del problema, en esta etapa se ha logrado afianzar nuestros conceptos vistos en clase hasta el momento y sobre todo la abstracción de las entidades.

En esta ocasión se ha podido analizar y discutir cada uno de las exigencias de la empresa, con el fin de adelantar lo más posible, teniendo en cuenta nuestro saber, además es probable que tengamos que corregir varios puntos de nuestro adelanto, pero es necesario pasar esta etapa ya que se practica y se afianza lo visto en clase y que a su vez ayuda a cumplir con nuestro objetivo del curso.

En este ejercicio nos hemos dado cuenta que el buen trabajo en equipo y la comunicación es sumamente importante, como también empezar a prever la manera en que vamos a desarrollar los objetivos del proyecto en los diferentes servidores, con sus respectivas características y diferencias para tener consistencia y orden en futuras entregas.

# RECOMENDACIONES

Para el desarrollo del ejercicio es importante repasar los conceptos vistos en clase para poder realizar las practicas, mucho mejor sin cometer errores que cuestan tiempo y trabajo en el futuro.

Además, es necesario tener muy presente las diferencias de la sintaxis dependiendo del cliente que se esté manejando, Por ejemplo: en el caso de Oracle se optó por usar secuencias en el campo de la llave primaria, donde en otros clientes se usó un identity, también es diferente la creación de la segunda vista, ya que a diferencia de los demás servidores es necesario enumerar los campos.

Se recomienda leer la documentación existente para cada uno de los servidores, estos nos permiten reconocer las diferencias entre las sentencias, la sintaxis y demás propias de cada herramienta, por esta razón, es indispensable considerarlas a la hora de crear una base de datos.

Por último es indispensable tener en cuenta los requerimientos básicos y técnicos del proyecto ya que de nada sirve realizar una base de datos que no cumple con las características requeridas.

# BIBLIOGRAFÍA

11/02/2019.Índices agrupados y no agrupados descritos. Documentacion de SQL.<https://docs.microsoft.com/es-es/sql/relational-databases/indexes/clustered-and-nonclustered-indexes-described?view=sql-server-ver15>

28/04/2020.Índice(Base de datos).Wikipedia.<https://es.wikipedia.org/wiki/%C3%8Dndice_(base_de_datos)>

Indice de bases de datos. Ecured. <https://www.ecured.cu/%C3%8Dndice_de_base_datos>

Robidoux G. 3/11/2011. index scan and table scan.MSSQLtips.<https://www.mssqltips.com/sqlservertutorial/277/index-scans-and-table-scans/>

Secuencias (create secuence-currval- nextval).<https://www.tutorialesprogramacionya.com/oracleya/temarios/descripcion.php?inicio=25&cod=193&punto=35#:~:text=Otro%20objeto%20de%20base%20de,sus%20valores%20no%20se%20repitan>.

14/03/2017.Vistas.Microsoft.<https://docs.microsoft.com/es-es/sql/relational-databases/views/views?view=sql-server-ver15>

12/102012.Definición de campo registro bases de datos.Estructura de datos.<http://estructura-davomtz92.blogspot.com/2012/10/definicion-de-campo-registro-base-de.html>

22/08/2019.Fundamentos de bases de datos- tablas,campos,registros. Akus.net.<https://disenowebakus.net/creando-bases-de-datos.php>

(02-03-2020).TCP(Transmission Control Protocol).Digital Guide ionos.https://www.ionos.es/digitalguide/servidores/know-how/que-es-tcp-transport-control-protocol/

Modelo cliente-servidor.Herramientas web para la enseñanza de protocolos de comunicación.https://neo.lcc.uma.es/evirtual/cdd/tutorial/aplicacion/cliente-servidor.html

Cunha.D(02-01-2020).Qué es un proxy.we live security.<https://www.welivesecurity.com/la-es/2020/01/02/que-es-proxy-para-que-sirve/>

Bases de datos NoSQL | Bases de datos orientadas a objetos.S,F.grapheverywhere.[https://www.grapheverywhere.com/bases-de-datos-orientadas-objeto](https://www.grapheverywhere.com/bases-de-datos-orientadas-objetos/)