

Segunda Entrega Proyecto Final

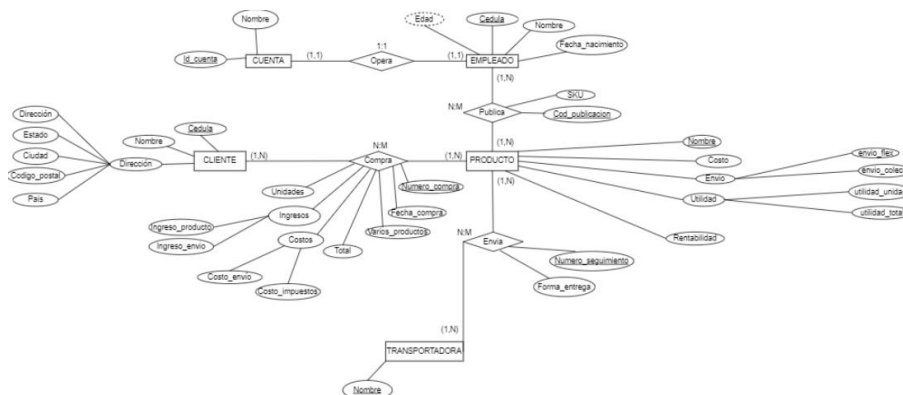
Integrantes: Juan Diego Buitrago y Santiago Gómez Corrales

Correos: juandi.buitrago@urosario.edu.co, santiago.gomezco@urosario.edu.co

Siguiendo el instructivo de la rúbrica para la segunda entrega procedimos a hacer las respectivas modificaciones identificadas en la primera entrega y completar las medidas para la segunda. Los cambios fueron los siguientes:

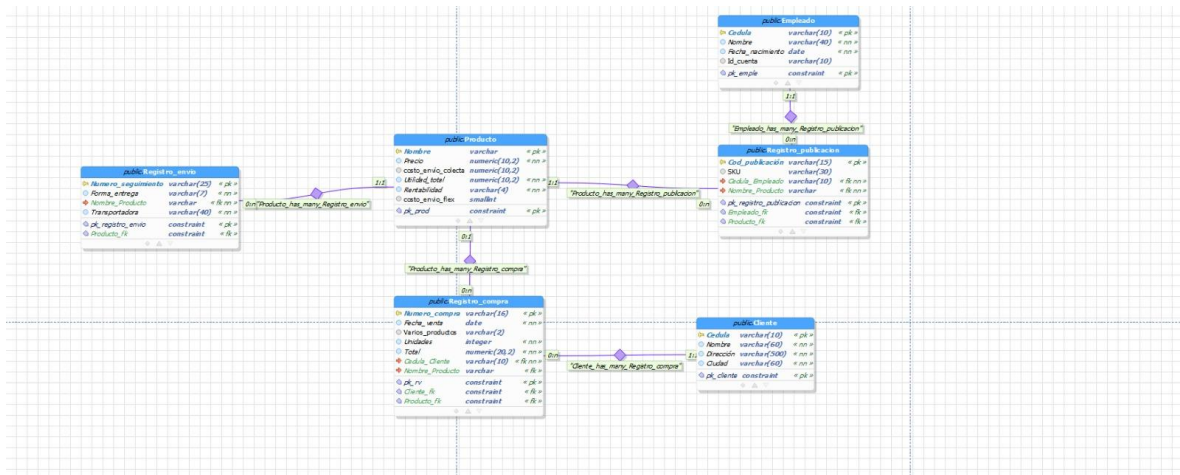
- En la parte inferior de este documento se encontrará el link del repositorio y el link del drive de nuestra base de datos. No nos fue posible subirla en github por errores que nos salían pero esperamos con ayuda de la profesora poder cargar esta al repositorio lo más pronto posible. Así que para no generar ninguna incomodidad proporcionamos el link del drive donde se puede ver perfectamente.
- Creación del repositorio
- Nos dimos cuenta que nuestra base de datos era bastante extensa, por lo que decidimos recortar todo al mes de abril, para así tener un óptimo proceso y desarrollo del proyecto
- Nos dimos cuenta que en general la idea de las tablas estaba bien, pero teníamos que modificar el tipo de datos de algunos, ya que o eran muy grandes o nos botaba error, por lo tanto, decidimos dejar esos de error como varchar para así poder continuar y para la próxima entrega tratar de tenerlos con el tipo de dato que teníamos presupuestado desde un inicio.
- Nos dimos cuenta que en el diagrama entidad-relación y en el diagrama relacional faltaba pulir algunas cosas, así que a continuación presentaremos ambos.

Diagrama Entidad Relación



Principalmente el cambio fue cambiar la ubicación de la Edad del empleado ya que gracias a las correcciones nos dimos cuenta del error que teníamos al ubicarlo junto con fecha de nacimiento, y el otro cambio fue quitar el NIT de la transportadora ya que nos dimos cuenta que podría generar conflictos a futuro entonces decidimos modificar eso de una vez.

Diagrama Relacional (Final)



Este es nuestro diagrama relacional final. Si bien es cierto tuvimos algunas complicaciones en esta entrega con la carga masiva de datos por algunos tipos de datos, nuestro objetivo es para la siguiente entrega tenerlo todo perfecto tal y como está dispuesto aquí, son errores pequeños por corregir con ayuda de la profesora para la siguiente entrega.

Carga de Datos:

Para la carga masiva de datos nosotros hicimos el siguiente proceso:

- Como nuestra base de datos se dividía por los días de cada mes lo que hicimos fue crear los distintos libros de excel y pasar día por día los datos, lo cual fue un proceso que nos tomó casi el 60% del tiempo empleado en esta entrega, pero finalmente pudimos hacerlo.
- Luego transformamos cada libro a csv pero teníamos unos errores así que tuvimos que buscar la manera de pasarlo a csv UTF-8 ya que en nuestros computadores no salía la opción para guardarlo así, así que empleamos otro computador para esto.
- Luego de esto fuimos a la monitoria de la profesora para verificar como era el comando para la copia masiva puesto que estábamos usando uno muy sencillo que podía botarnos muchos errores, y con esta corrección entendimos a la perfección la copia masiva de datos con csv y el comando copy.
- Finalmente pudimos crear todas las tablas con sus datos en PGAdmin4 en su formato SQL.
- Finalmente, con la ayuda de la librería psycopg2 pudimos hacer lo de Python y cargamos todo al repositorio.

Análisis de Escenarios:

Considerando los pasos que hemos seguido para llegar hasta este punto tanto como las pruebas como las investigaciones llevadas a cabo entorno a la base de datos, planteamos las siguientes perspectivas de análisis. Con esto buscamos formular la propuesta de explorar y evaluar diversos escenarios de análisis.

Con estos análisis buscamos comprender mejor el funcionamiento y explicarlo de manera óptima, para así ver hasta dónde podría llegar su alcance. Buscamos una visión completa y contextualizada del proyecto que estamos llevando a cabo con la base de datos de la empresa SD Grupo Empresarial.

- Analizar la eficiencia de cada trabajador de SD Grupo Empresarial para así poder evaluar sus estadísticas laborales durante el día, mes o año que se requiera. Esto ayudará a la empresa estadísticamente y le permitirá identificar si sus empleados están cumpliendo con las metas presupuestadas ya que para esta empresa puede ser complicado medir esto.
- Llevar a cabo un análisis de los productos más pedidos. Esto nos proporcionará una estadística que nos permitirá ver tanto el producto más solicitado como el género del producto, ósea, para que área en específica se emplea este producto y así poder mostrar más de esto por medio de la publicidad propuesta por la empresa para el crecimiento exponencial de la misma.
- Examinar la disposición geográfica de los compradores. Con esto queremos evaluar las partes en donde más se piden cosas por medio de la empresa, para así mantener un fuerte estándar sobre estas zonas y en las más débiles proporcionar ideas para que haya un crecimiento equitativo y exponencial.
- Evaluar los rangos de precio que hay entre los distintos productos. Con esto podríamos evaluar en promedio el presupuesto que se gasta una persona en un pedido con la empresa, con esto podríamos dar más publicidad y conseguir más productos de dicha índole para llamar la atención de los compradores y así generar más demanda en dicho rango de precio.
- Investigar la preferencia de los compradores con respecto a las transportadoras. Con esto nos podríamos hacer una idea de si enserio vale la pena mantener o crear nuevas relaciones con empresas transportadoras o si en contrario es mejor quedarse con la base y trabajar solo con esa así ahorrando capital puesto que los usuarios no emplean las otras. Esto es un escenario de lo que buscamos investigar con respecto a las transportadoras.

Link Repositorio: <https://github.com/Santhiago04/proyectoDatos>

Link Base de Datos: [Base de Datos DRIVE](#)