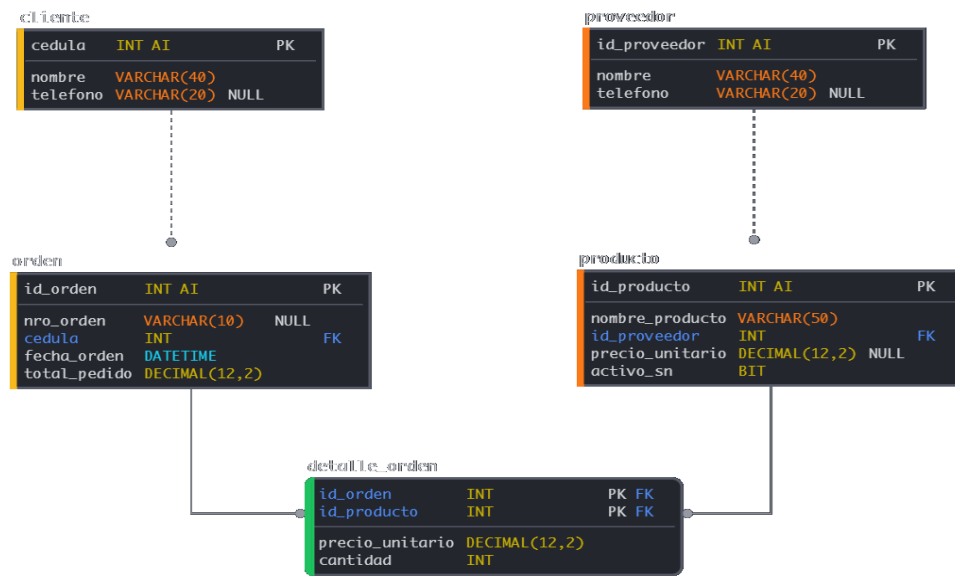


## PRUEBA ANALISTA DE DATOS – SEGUIMIENTO

**Objetivo:** Evaluar las capacidades técnicas de los candidatos a través de un caso de uso en el cual se aplique un proceso completo enfocado en datos, dividido en 3 fases.

### Fase 1: Extracción y Combinación de Datos

**Descripción:** En esta fase, el candidato deberá demostrar su habilidad para extraer datos en SQL de diversas fuentes y combinarlos para generar un modelo de datos coherente y utilizable. Tenga en cuenta el diagrama entidad relación para la construcción de sus respuestas.



De acuerdo con el diagrama, escribir las consultas SQL que arrojen los siguientes resultados:

1. Número total de órdenes registradas.
2. Número de clientes que han realizado órdenes entre el 01-01-2021 y la fecha actual.
3. Listado total de clientes con la cantidad total de órdenes realizadas (conteo), ordenando de mayor a menor nro. de órdenes.
4. Detalle completo (datos del cliente, fecha, nombre producto, cantidad) del pedido cuyo monto fue el más grande (en valor, no en unidades) en el año 2020.
5. Valor total vendido por mes y año.
6. Para el cliente con cédula 123456, especificar para cada producto, el número de veces que lo ha comprado y el valor total gastado en dicho producto. Ordenar el resultado de mayor a menor.
7. Si necesitas actualizar una tabla histórica con los datos del último mes, y en este nuevo mes has incluido una nueva columna para la ciudad del cliente, ¿qué proceso seguirías para evitar conflictos por diferencia de dimensiones, considerando que no tienes acceso a los comandos ADD COLUMN o ALTER TABLE?

**Nota:** Se evaluarán buenas prácticas de SQL.

## Fase 2: Análisis de Información

**Descripción:** El candidato deberá utilizar los datos almacenados en el archivo "*Productividad.xlsx*" para generar un análisis con el cual logre identificar cual es el cargo más conveniente contratar enfocado en incrementar la productividad.

1. Tenga en cuenta que el costo asociado de cada cargo es 100% tiempo completo, 50% medio tiempo, 40% fines de semana.

## Fase 3: Clasificación de los puntos de venta en segmentos

**Descripción:** El candidato deberá crear un modelo utilizando técnicas de aprendizaje no supervisado para identificar patrones y generar estrategias basadas en los datos utilizando el data set "*Segmentacion\_CATT.xlsx*".

## Fase 4: Presentación de Resultados

**Descripción:** El candidato presentará los resultados obtenidos, explicando la selección de variables, el tipo de modelo o metodología utilizada, y los hallazgos más importantes que le permitan justificar que cargo de ejecutivos deberíamos contratar y en qué puntos.

## Resultados esperados:

**Descripción:** El candidato deberá presentar conclusiones generales combinando las diferentes fases de la prueba (2, 3 y 4) con el objetivo de incrementar los niveles de productividad de la compañía basadas en los resultados obtenidos y sugerir próximos pasos para mejorar o expandir el análisis.

## Aspectos para tener en cuenta:

1. La solución técnica debe crearse en un repositorio de GitHub (público), el cual debe tener un documento README en el cual explique el desarrollo que dio a cada ejercicio.
2. La presentación de resultados puede ir acompañada de un PDF, Power Point e incluso un Dashboard.