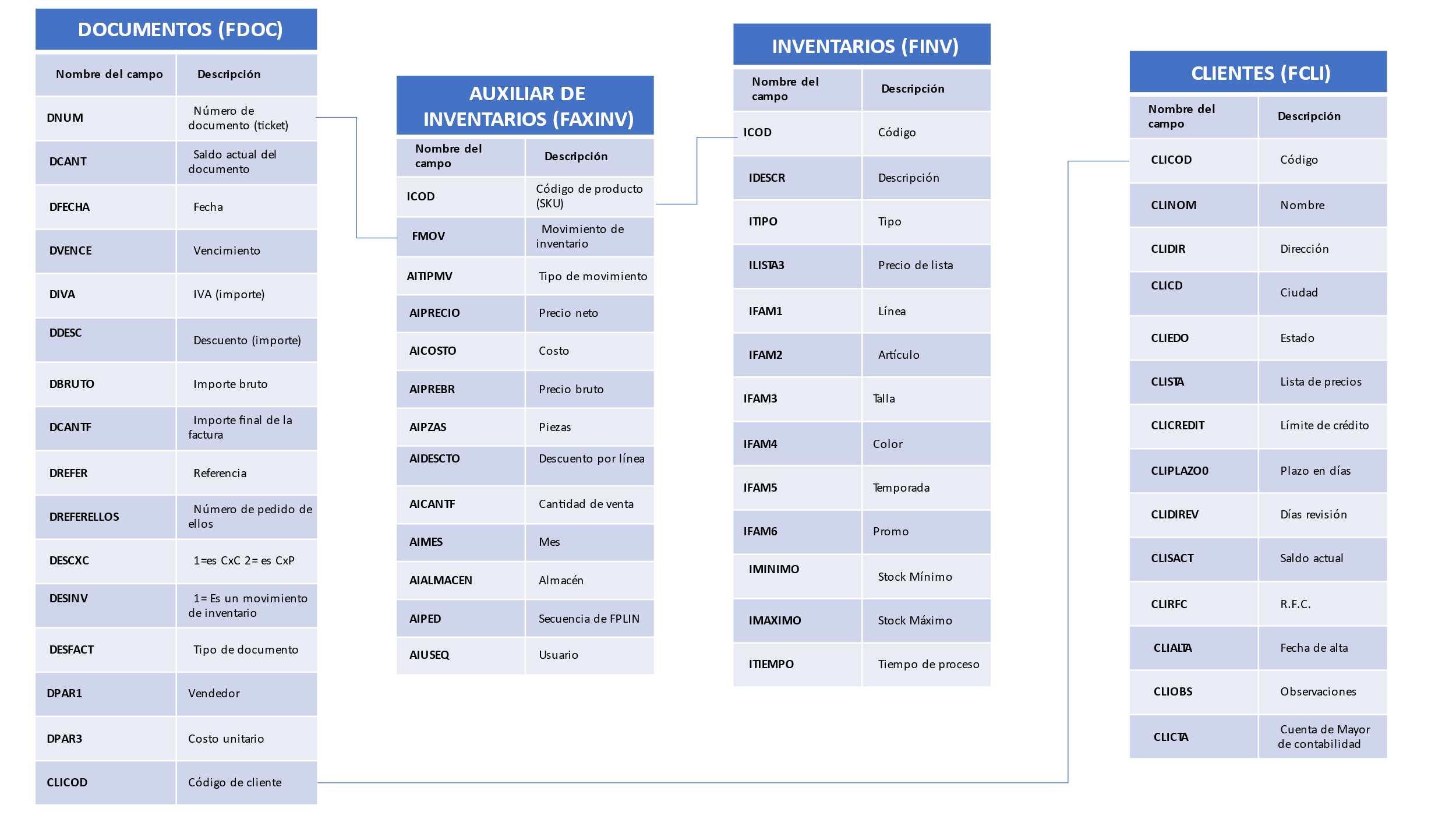
Juan Daniel Almeyda Calapiz

1. **Construcción de query en SQL**

**Instrucciones:** Con el diagrama de base de datos compartido, crea un query para obtener la información de ventas del cliente, incluyendo los siguientes campos:

* # de ticket
* Fecha
* Monto (DCANT) con y sin IVA
* Vendedor
* Cliente
* Productos comprados (SKU)
* Descripción
* Cantidad (unidades)
* Almacén que vendió
* Precio de lista
* Talla
* Color
* Temporada



**Entregable:** Código escrito con la consulta solicitada. Puedes asumir que las tablas ya están creadas en la base de datos.

SELECT d.DNUM No\_Ticket,

d.DFECHA Fecha,

d.DCANT Monto,

d.DCANT + (d.DCANT\*0.16) Monto\_IVA,

d.DPAR1 Vendedor,

c.CLINOM Cliente,

i.ICOD Producto\_SKU,

i.IDESCR Descripcion,

ai.AICANTIF Cantidad,

ai.AIALMACEN Almacen,

i.ILISTA3 Precio\_lista,

i.IFAM3 Talla,

i.IFAM4 Color,

i.IFAM5 Temporada

FROM FDOC d

JOIN FCLI c ON d.CLICOD = c.CLICOD

JOIN FAXINV ai ON d.DNUM = ai.FMOV

JOIN FINV i ON ai.ICOD = i.ICOD

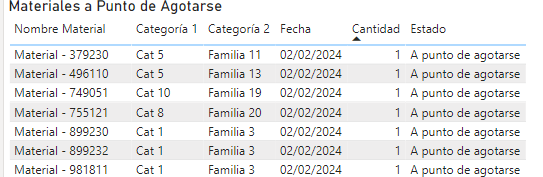
WHERE c.CLICOD = 'Codigo'

1. **Desarrollo de visualizaciones en Business Intelligence**

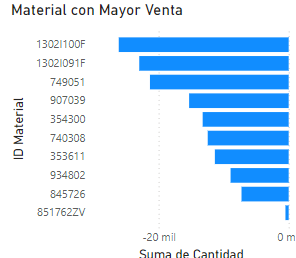
**Instrucciones:** Con el archivo “Fuente para dashboard”, desarrolla un dashboard que permita responder la mayor cantidad de preguntas de la siguiente lista, para dar una recomendación al cliente acerca de su operación.

**Preguntas:**

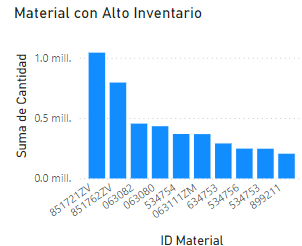
1. ¿Qué materiales me voy a quedar sin inventario y en qué fecha me quedo sin inventario?

Los principales materiales que estan a punto de quedarse sin inventario son:  


1. ¿Cuáles productos son los más importantes en venta?

Los principales productos con mayor venta son:  


1. ¿Cuáles productos tengo riesgo de quedarme con mucho inventario al final del año?

Los productos con riegos de quedarse con mucho inventario son:  


1. ¿Hay mucha variación entre la venta de años pasados y el forecast?

El año que presenta mayor variacíon entre las ventas y el forecast es el 2023, ya que se tiene mayor historial de ventas, para el año 2021 se presentan menos ventas

1. ¿Cuál sería la cobertura del inventario contra la versión 1 del forecast y con la versión 2 del forecast?

Existe una cobertura del 0.01 % entre el forecast versión 1 a la versión 2

1. Después de haber presentado esta primera versión del dashboard, ¿qué pasos recomendarías seguir para liberarlo para el uso en el día a día por parte del cliente?

Verificar que la información corresponda correctamente realizando pruebas con los datos que se tienen, mostrar cómo funciona cada visualización e incluso implementar mejoras en cuanto a diseño.

1. ¿Cómo esperarías que fuera la interacción entre IBSO y el cliente por los siguientes 12 meses?

Esperaría una interacción constante, ya sea de forma diaria para verificar que los datos se estén actualizando correctamente, mantener comunicación en cuanto a futuros cambios y mantenimiento y para las capacitaciones que se sigan necesitando.

**Entregable:** Dashboard (En Tableau / Power BI / Qlik) con las tablas, gráficas y/o visualizaciones que justifiquen los resultados y recomendaciones propuestas. El dashboard debería seguir una estructura lógica y deberá permitir al usuario entenderlo sin mucha capacitación.

**Referencias para estructura del dashboard:**

<https://medium.com/@mokkup/8-essential-dashboard-design-principles-for-effective-data-visualization-40653c5fd135>

Tableau: <https://data.ucop.edu/support-training/tableau-files/building_effective_dashboards.pdf>

1. **Ejercicios – Python**

**Instrucciones:** Con las preguntas mostradas, trata de desarrollar los códigos para responder el mayor número de ejercicios usando Python.

**Entregable:** Códigos en Python mostrando los resultados de los ejercicios.

**Ejercicio 1:**

1. Crea un diccionario llamado datos\_climaticos donde cada clave es el nombre de una ciudad y el valor es una lista de temperaturas (en grados Celsius) registradas a lo largo de una semana.
2. Para cada ciudad, calcula la temperatura promedio, la temperatura máxima y la mínima de la semana.
3. Determina cuál fue la ciudad con la temperatura promedio más alta y la más baja durante la semana.

**Ejercicio 2:**

Asigna a cada letra minúscula un valor, desde 1 para la 'a' hasta 26 para la 'z'.

Crea una función que pida al usuario una cadena de letras minúsculas y responde la suma de los valores de las letras en la cadena. (Ejemplo: hola = 8 + 15 + 12 + 1 = 36).

Si el usuario te da un número o una letra mayúscula, pídele que lo cambie (Input: Hola. Output: Cambia a minúscula la letra “H” en la posición 1. Input: int2. Output: Cambia el número en la posición 4 por una letra minúscula).

**Ejercicio 3:**

1. Extraer la información del csv **Prueba\_Promociones**

2. ⁠Generar un código donde el usuario pueda ingresar las siguientes variables

- fecha inicio (convertir a datetime)

- ⁠fecha fin (convertir a datetime)

- ⁠categoría (validar que sea string)

- ⁠uso (validar que sea string)

- ⁠sku (no permitir al usuario avanzar si no ingreso un valor string en el campo de SKU)

- ⁠% (validar que sea número decimal)

- ⁠inventario inicial (validar que sea entero)

3. Generar una nueva columna donde se coloque el # de semana correspondiente de la fecha

4. ⁠Con él % ingresado por el usuario, impactar ese porcentaje como crecimiento (columna de piezas) para todos los datos que cumplan con las siguientes condiciones:

- estén dentro del rango de fecha seleccionado. Si el usuario no coloca fecha fin entonces desde fecha inicio hasta el final, si el usuario no coloca fecha inicio entonces desde la fecha fin hasta el principio y si no coloca ninguna, a todas las fechas disponibles

- ⁠la columna **modelo** sea diferente a “real”

- ⁠la columna **uso** sea igual al valor ingresado por el usuario, en caso de que esté vacío ese campo, la columna categoría sea igual al valor ingresado por el usuario. En caso de que ambas vengan vacías aplicar a todo

5. Una vez aplicado el incremento %, generar un nuevo dataframe en donde solo venga información del SKU seleccionado y se pueda visualizar el consumo de inventario, es decir, el inventario inicial colocado por el usuario será la cantidad inicial el 01-01-2024 y de ahí se tendrá que ir consumiendo (restando) la venta proyectada (**modelo** diferente de real). Se deberá imprimir la primera fecha en la que se vuelve negativo el inventario

**Ejercicio 4: Recomendador de libros**

Fuente: <https://developer.nytimes.com/docs/books-product/1/overview>

**Elementos a utilizar:**

1. Crea una cuenta de Developer en la página de New York Times para tener un API key (<https://developer.nytimes.com/>)
2. Usa el “Books API”.

**¿Qué debe poder hacer tu código?**

1. Pedirle al usuario decidir qué lista de “Best Sellers” quiere consultar.
2. Poder escoger si quiere ver los “Best Sellers” actuales o de alguna fecha en específica.
3. Poder escoger un precio específico del libro que quiere adquirir.
4. Poder escoger un rango de edades dirigido para el libro.

**Resultado:**

1. Poder verlo de manera estructurada, poniendo la información clave del libro que estás recomendado para la información que te dio el usuario.
2. Poder acceder rápidamente a la reseña generada por el NYT acerca del libro recomendado.
3. Para los Best Sellers actuales, decirle al usuario dónde lo puede comprar.
4. Poder mostrar toda esta información en una aplicación / pantalla / interfaz gráfica amigable para que el usuario pueda interactuar con los resultados.
5. **Examen técnico**

Explica con el mayor detalle las respuestas a estas preguntas (45 minutos).

1. La información del ERP del cliente vive en un data center controlado por un proveedor, con reglas que no permiten adicionar campos, columnas, construir tablas o extraer información. ¿Qué pasos harías para pasar de la información en el ERP, cruzar y transformarla, y mandarla a una herramienta de Business Intelligence con actualizaciones automáticas?

Realizaría un proceso de ETL, utilizando SQL Server Integration Services (SSIS), solicitando el acceso a las API’s del ERP dependiendo los permisos para extraer los datos necesarios ya sea de manera semanal o por día, también utilizando jobs que nos permitan ejecutar las extracciones de manera automática, por medio del servicio de Integracion transformaria los daros de acuerdo a las necesidades para después almacenarlos en Data Warehouse propio o una base de datos especifica dentro del motor de SQL Server y así poder manipular los datos y/o analizarlos, conectaría la fuente de datos a través de la herramienta Power BI, creando el reporte respecto de manera eficaz y entendible para los usuarios que tengan acceso a él, por ultimo realizaría la configuración necesaria, Power Bi utiliza un archivo de configuración que se tiene que instalar dentro del equipo necesario que permitirá la conexión de la fuente de datos a los diferentes reportes, así se realizara la actualización de forma automática cada vez que los nuevos datos sean extraídos, estableciendo diferentes periodos de automatización.

1. El cliente de la pregunta 1 está evaluando adquirir un servidor físico o en la nube para poder alojar ahí un espejo de la información de su ERP. ¿Qué recomendarías en cuanto a costo o inversión y el rendimiento esperado dadas las características de cada una de las opciones?

Para esto es importante analizar las ventajas y desventajas que ofrece cada tipo de servidor, un servidor físico ofrece control total con respecto para lo que lo desees, pero la inversión que se realizara va a ser alta, esta será de manera inicial dependiendo todos los componentes que se necesiten y a su vez conforme vayan aumentando las necesidades y el mantenimiento será necesario cubrir más gastos, un servidor en la nube ofrece baja inversión y cubriendo solo que se necesita, además que el proveedor del servidor realizara el mantenimiento y actualizaciones necesarias, las desventajas es que es indispensable el uso de internet y el rendimiento depende de la conexión, se tiene menos control en cuanto a seguridad y configuración, en cuanto a la necesidad del cliente yo recomendaría el uso de un servidor en la nube ya que le permitirá almacenar la información de su ERP, así evitara altos gastos y podrá mantener ese respaldo para obtenerlo cuando sea necesario

1. Un cliente solamente cuenta con un servidor de producción donde viven los datos, ¿qué sugerirías para no afectar el desempeño de esta al conectar BI?
   1. ¿Cuáles serían las implicaciones (tiempo, conocimiento, administración, costo) de esta sugerencia?

Implementaría un Data Warehouse donde almacenaría los datos de producción, extrayéndolos de forma recurrente, así la herramienta de BI se conectaría a este tipo de almacén en vez de hacerlo directo de producción, en cuanto a tiempo depende la configuración del Servicios con respecto a la cantidad que se manejara, es necesario administrarlo correctamente para revisar que los datos se estén actualizando correctamente, los costos implican las licencias de software, así como los gastos de mantenimiento.

1. Un query que escribiste regresa 10,000 filas únicamente, pero su ejecución toma entre 3 y 4 horas. Por lo mismo, muchas veces no llegan los resultados completos a BI o se pierde la conexión mientras esto ocurre. ¿Qué harías?

Se necesita realizar un analisis al Query de consulta, revisando si habrá algún inconveniente en los índices de las tablas o si habrá joins que sean los que hagan que sea lento, si es posible reducir los joins que involucran cada tabla, revisar que las cláusulas where estén funcionando correctamente, además utilizaría una división en la consulta, ejecutarla en diferentes partes para mejorar el rendimiento de la consulta, esta división se puede manejar por periodos de tiempo si asi lo permite.

1. Te busca un cliente para decirte que los resultados que ve en los tableros no le hacen sentido (ayer si mostraba resultados correctos). ¿Qué pasos harías para darle una respuesta al cliente?

Es importante conocer cuáles son los inconvenientes que presenta el cliente, averiguar si hubo cambios en los datos o se agregaron nuevas cosas, de ser así aplicarlos inmediatamente, después verificar si se están extrayendo los datos correctamente, los servicios de Integración que no presenten algún problema, así como verificar si estos no se habrá presentado algún cambio y de ser necesario aplicar actualizaciones de software y por último verificar la configuración del tablero, verificando que todos procesos, métricas, cálculos se estén ejecutando correctamente.

1. ¿Qué mejores prácticas sugerirías para garantizar la seguridad de la información del cliente, ya sea en un servidor físico o en la nube?

Es importante mantener una administración de monitoreo en la información verificando cualquier inconveniente que se presente, mantener seguridad en las conexiones de internet, así como seguridad de forma física, ya sea de servidores, archivos, etc., lo primordial es contar con respaldos que sean de forma regular para poder realizar recuperaciones cuando sea necesario y pues como en todos lados es importante tener normas de seguridad y aplicar auditorías.

1. El director comercial de una empresa nos pide un proyecto para poder visualizar información en un tablero de BI con actualizaciones cada 15 minutos con información proveniente de distintas fuentes:
   1. ERP (SQL Server)
   2. CRM (API)
   3. Catálogos de información (Excel)
   4. Google Analytics
   5. Google Ads
   6. Shopify
   7. Facebook Ads

Prepara una propuesta técnica que incluya tiempos, costos y herramientas a utilizar para entregarle al director comercial este tablero. En caso de que haya información que necesites para hacer esta propuesta y no la tengas, genera supuestos para los que sería valida la propuesta.

De acuerdo a información recabada se presentan los siguientes costos:  
Licencias: $10,000/mes  
Desarrollo: $100,000  
Costos de Mantenimiento: $25,000/mes

Herramientas ETL: SQL Server Integration Services (SSIS) contiene los conectores necesarios a los diferentes tipos de API y fuentes donde se maneja la información.  
SQL Server: motor de base de datos, así como creación de Data Warehouse.  
Herramientas BI: Power BI ofrece opciones avanzadas de visualización.  
  
Tiempos:  
Recopilación de las necesidades del cliente: 1 semana  
Configuración de ETL: Para la Integración de las diferentes fuentes 2 semanas  
Configuración Base de datos: 1 semana  
Desarrollo del Reporte BI: 2 semanas  
Pruebas: 1 semana  
Capacitación: 1 semana  
Implementación en productivo: 1 semana

Toda implementación propuesta llega a ser adaptable a los requerimientos necesarios, ya sea incrementos en cantidad de datos o cambios en algunas fuentes de información, se implementan medidas de seguridad, respaldos y confidencialidad de la información, se deberá entregar un tablero funcional y entendible, coordinando también revisiones contantes para que el proyecto marche conforme a lo establecido.