

**UNIVERSIDAD DE EL SALVADOR FACULTAD DE INGENIERÍA Y
ARQUITECTURA ESCUELA DE INGENIERÍA DE SISTEMAS
INFORMÁTICOS.**



CURSO DE ESPECIALIZACIÓN INGENIERÍA DE DATOS
Ing. Rene Fabricio Quintanilla Gomez

Primera fase: Diseño de un modelo multidimensional para soportar un proceso de negocio seleccionado

Tema: Construcción de data warehouse para análisis de inventarios para las sociedades de grupo probe.

Grupo: 01

carnet	Apellidos	Nombres
BD03006	Beltrán Domínguez	José Eduardo
EQ 17001	Escobar Quintanilla	Francisco Josué
PG14048	Palacios Gonzáles	Gonzalo de Jesús
RP18045	Reyes Perdido	German Oswaldo

San Salvador, 26 de junio de 2025

Tabla de contenido

I. Introducción.....	1
II. Sistema de objetivos.....	2
III. Alcances.....	3
IV. Situación Actual.....	4
A. Antecedentes.....	4
B. Descripción del problema.....	7
C. Planteamiento del problema.....	8
V. Hipótesis.....	9
A. Hipótesis nula.....	9
B. Hipótesis alterna.....	9
VI. Justificación.....	10
VII. Metodología de trabajo.....	12
VIII. Descripción de la propuesta de solución.....	16
IX. Tecnologías a utilizar.....	17
X. Diagrama arquitectónico de la solución.....	19
XI. Descripción de cada componente de la solución.....	20
XII. Cronograma de actividades.....	23
XIII. Presupuesto.....	25
XIV. Anexos.....	26

I. Introducción

En este documento abordaremos el Análisis de Inventarios para la Empresa de grupo Probe, ellos son una corporación salvadoreña con más de 45 años de trayectoria en el sector de la belleza, estética y distribución profesional de productos cosméticos. La empresa está compuesta por diversas Sociedades como Vidals, Sento Luxury Spa & Salón, TBS Beauty Supply Disprobe / Intrade.

Uno de los aspectos más importantes de Grupo Probé es que invierte activamente en educación técnica, operando centros de formación que han capacitado a miles de profesionales de la belleza, promoviendo así el emprendimiento y la profesionalización del sector.

Ellos cuentan con una solución de inteligencia de negocios basada en cubos OLAP implementados sobre infraestructura Microsoft SQL Server, cuya finalidad es consolidar y analizar información crítica principalmente relacionada a inventarios de las distintas sociedades operativas

El análisis detallado del inventario les permitirá a probe gestionar de forma más ordenada y con rapidez el crecimiento de las operaciones y el volumen de información que día con día aumentan.

Permitirá mejorar el manejo de las vistas en tiempo real, carga de información de forma más eficiente, replicar procesos en la demás sociedad que las conforman y el manejo de volumen de datos con un mejor control y resguardo de cada uno de los productos que manejan dentro de las sociedades que la conforman desde la data hasta los productos en existencias generando así reportes en tiempo real para la toma de decisiones y así hacer uso del activo más valioso que es la información.

Permitirá que grupo probe gestione inventarios de forma más rápida integrando así las distintas unidades de negocio que la conforman, consolidando el inventario en tiempo real y así contar con la información de primera mano que hoy en día es su mayor reto.

II. Sistema de objetivos

General

- Construir un Data Warehouse que permite centralizar, almacenar y analizar de manera eficiente todos los datos de inventarios de las sociedades de Grupo Probe, con el fin de generar reportes BI y métricas que mejoren la toma de decisiones en la gestión de inventarios.

Específicos

- Realizar un levantamiento y análisis de los requisitos solicitados y de los procesos que actualmente se manejan de inventarios en las sociedades de Grupo Probe, identificando todas las posibles fuentes de datos y definiendo los indicadores o reportes a generar.
- Diseñar la arquitectura conceptual y lógica del Data Warehouse, incluyendo el modelo dimensional para inventarios, las estrategias elegidas de ETL y toda la infraestructura necesaria para su implementación
- Utilizar la infraestructura del Data Warehouse y desarrollar los procesos de extracción, transformación y carga (ETL) conforme a el diseño previamente aprobado.

III. Alcances

El proyecto se basa principalmente en la construcción de un data warehouse de varios procesos de negocio de un rubro o un ámbito específico, en este caso son los inventarios determinados por la empresa probe.

El data warehouse busca sustituir los cubos OLAP con el paso del tiempo, debido a que no puede sustituirse inmediatamente por la cantidad de información histórica que se utiliza actualmente. El data warehouse se sitúa como una solución que sea mantenible y con mejor rendimiento del que ya provee el cubo OLAP, debido a que este deja de ser una solución a la generación de información gerencial por la alta cantidad de tiempo que requiere disponibilizar la información para los usuarios.

Se realizará un data warehouse que de solución a la analítica de procesos de negocio tales como:

- Ingreso de productos a una bodega
- Devoluciones de productos
- Requisiciones o salidas de productos a una bodega
- Transferencias entre bodegas de una empresa.

Los procesos mencionados anteriormente son un ejemplo de posibles procesos de negocios orientados a inventarios, estos deben ser delimitados por la empresa Probé, la cual es el principal actor en determinar los procesos de negocio que necesitan esta herramienta de análisis de datos.

Posteriormente, se deberá generar la herramienta de recolección, transformación y carga de datos y toda la documentación correspondiente al data warehouse, tal como diagramas de base de datos, modelo físico, diccionario de datos, entre otros.

El proyecto no contempla la implementación del data warehouse en un tablero que pueda ser visualizado en un ambiente de producción, sin embargo, se realizarán visualizaciones previas para que la empresa pueda visualizar el seguimiento y el progreso del proyecto.

Los principales entregables del proyecto son:

- Script para generar el data warehouse
- Los ETL utilizados para extraer la información.
- Documentación de los data warehouse y los ETL

IV. Situación Actual

A. Antecedentes

Grupo Probe es una corporación salvadoreña con más de 45 años de trayectoria en el sector de la belleza, estética y distribución profesional de productos cosméticos. Fundada en 1978, se ha consolidado como una de las empresas líderes en El Salvador, con un modelo de negocio integral que abarca servicios, formación, distribución y venta de productos de belleza.

La empresa está compuesta por diversas Sociedades:

- **Vidals:** La cadena de salones de belleza más grande del país, especializada en estilismo, cuidado capilar y estética tanto para hombres como para mujeres. Vidals cuenta con múltiples sucursales en todo el territorio nacional.
- **Sento Luxury Spa & Salon:** Un espacio de bienestar y relajación de alto nivel, que ofrece más de 350 servicios personalizados, combinando tecnología de punta con un enfoque integral de cuidado personal.
- **TBS Beauty Supply:** Red de tiendas enfocada en la venta de productos de belleza y cosmética profesional. Atiende tanto a consumidores finales como a profesionales del sector.
- **Disprobe / Intrade:** Las divisiones encargadas de la distribución mayorista de reconocidas marcas internacionales del rubro cosmético, abasteciendo a salones, spas y tiendas especializadas en todo el país.

Además, Grupo Probe invierte activamente en educación técnica, operando centros de formación que han capacitado a miles de profesionales de la belleza, promoviendo así el emprendimiento y la profesionalización del sector.

Con más de 480 colaboradores y un compromiso continuo con la innovación, la excelencia en el servicio y la transformación digital, Grupo Probé ha evolucionado hacia una empresa con visión estratégica, implementando soluciones tecnológicas avanzadas como sistemas ERP, POS, y plataformas de facturación electrónica.

En cuanto a los trabajos realizados en la empresa a nivel de ingeniería de datos se cuenta con una solución de inteligencia de negocios basada en cubos OLAP implementados sobre infraestructura Microsoft SQL Server, cuya finalidad es consolidar y analizar información crítica principalmente relacionada a inventarios de las distintas sociedades operativas: Vidals, Sento, TBS, Disprobe e Intrade. Esta solución está en funcionamiento y ha sido útil para habilitar el análisis multidimensional de datos a través de herramientas como Excel o interfaces web internas desplegadas en la intranet corporativa.

El modelo implementado se alimenta mediante vistas SQL que sirven como capa de abstracción entre los sistemas fuente (OVST, POS Softland, Softland ERP) evitando procesos ETL complejos. Estas vistas, adaptadas para cada sociedad, transforman los datos operacionales en estructuras compatibles con el cubo OLAP construido en SQL Server Analysis Services (SSAS).

No obstante, este enfoque aunque funcional se ha vuelto insostenible frente al crecimiento de la operación y el volumen de información.

puntos en la problemática técnica en la arquitectura actual:

1. Rendimiento limitado por el uso exclusivo de vistas SQL

El modelo actual se basa en un conjunto de vistas complejas que actúan como capa de integración entre los sistemas fuente y el cubo OLAP. Estas vistas realizan operaciones pesadas como Joins entre múltiples tablas, conversiones de datos, agregaciones y validaciones condicionales, lo cual carga en tiempo real al motor de base de datos.

Cada vez que se consulta o actualiza el cubo, estas vistas se ejecutan desde cero, procesando volúmenes crecientes de datos históricos. Este enfoque no solo genera cuellos de botella que afectan los tiempos de carga del cubo y las consultas en Excel, sino que compite directamente por recursos con el sistema ERP Softland, al estar ambos sobre la misma instancia de SQL Server. Como consecuencia, se producen ralentizaciones generales que pueden impactar operaciones críticas del negocio en horarios productivos.

2. Carga total (full load) obligatoria del cubo OLAP

Debido a que no existe un mecanismo de carga incremental (como Change Data Capture o marcas de modificación), el cubo debe reprocesarse por completo en cada ciclo de actualización. Esto implica volver a cargar todos los datos históricos de cada sociedad, sin distinguir entre lo nuevo y lo ya consolidado.

Esta metodología es ineficiente y consume tiempo y recursos excesivos, especialmente en un contexto donde el volumen de datos crece constantemente. Como resultado, se ha tenido que limitar la frecuencia de actualización del cubo, restringiéndola a procesos nocturnos o de fin de semana, para evitar interferencias con la operación regular. Esto introduce desfases entre el estado real del inventario y el que se muestra en los reportes, afectando decisiones de reposición, Ingresos y consumo.

3. Fragmentación lógica entre sociedades

Cada sociedad Vidals, Sento, TBS, Disprove, Intrade cuenta con lógicas de negocio propias y estructuras de datos distintas, lo que ha llevado a construir vistas separadas y personalizadas para cada una. Si bien esto permitió flexibilizar el desarrollo inicial, ha derivado en una fragmentación estructural que impide una visión homogénea del inventario.

No existe una capa semántica común ni un modelo de datos centralizado que garantice la coherencia de los indicadores. Esto provoca discrepancias entre métricas similares, como “consumo interno”, “ajustes por pérdida” o “stock disponible”, dependiendo de la fuente. Además, dificulta el mantenimiento técnico, ya que cada modificación debe aplicarse manualmente en múltiples lugares, con riesgo de errores o desalineación entre sociedades.

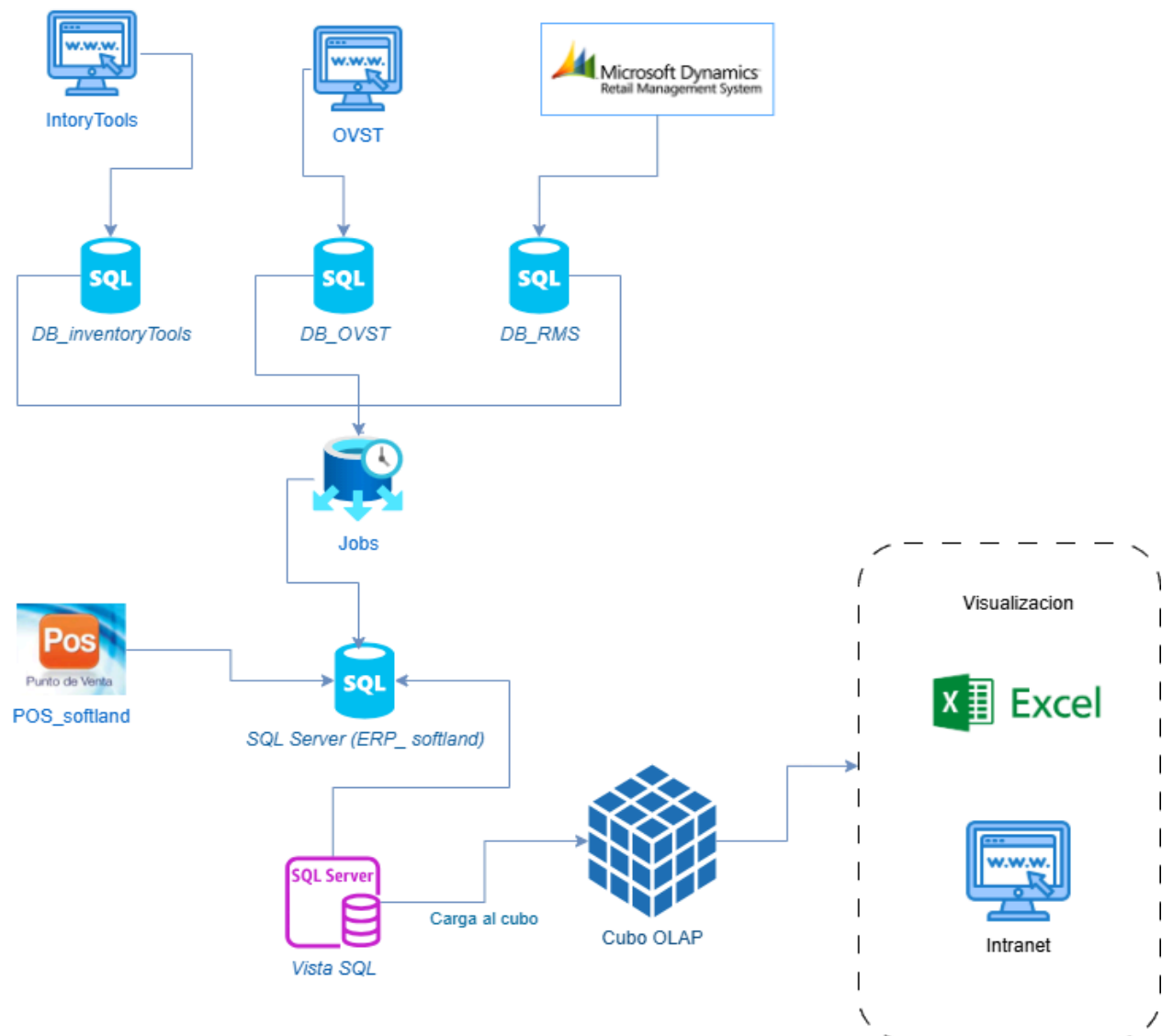
4. Dificultades para escalar la solución

La arquitectura actual fue útil en fases iniciales del proyecto, pero no fue diseñada para escalar. A medida que aumentan los volúmenes de datos, el sistema carece de componentes fundamentales para soportar el crecimiento:

- No hay particionamiento de datos por fechas o unidades de negocio.
- La capacidad de almacenamiento no está optimizada para consultas analíticas pesadas.
- Falta una gobernanza de datos que regule definiciones, accesos y estructuras.

Estas carencias impiden incorporar nuevos indicadores, sociedades o dimensiones sin comprometer el rendimiento general. Además, obstaculizan la incorporación de nuevas tecnologías analíticas (como modelos predictivos o monitoreo en tiempo real), dejando a la empresa con un sistema rígido, reactivo y con riesgo de obsolescencia técnica.

Este diagrama representa el estado actualmente la infraestructura:



B. Descripción del problema

Con lo antes mencionado la problemática que tiene Grupo Probe es la gestión de sus datos, la cual se deriva de la fragmentación de sus sistemas de información y la falta de integración entre sus distintas unidades de negocio. Siendo una empresa con operaciones distribuidas en múltiples Sociedades Vidals, Sento, TBS Beauty Supply como se conocen comercialmente, y su matriz distribuidora Disprobe/Intrade, la consolidación y control de datos en tiempo real se ha vuelto uno de sus principales retos operativos.

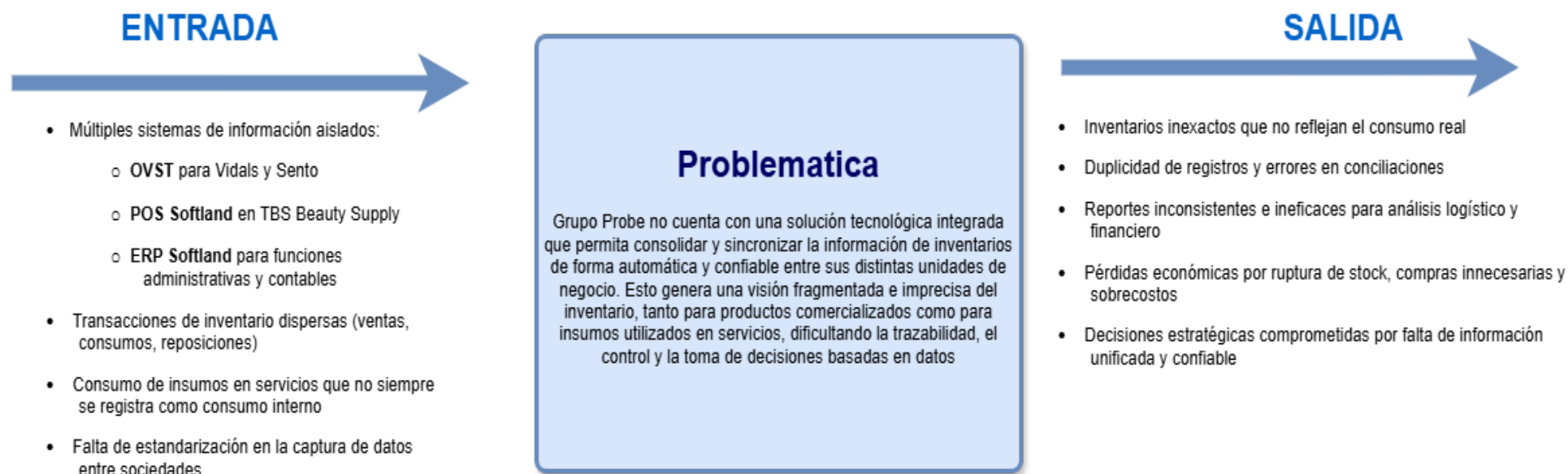
Históricamente, Vidals y Sento trabajaron con el sistema Microsoft Dynamics Retail Management System (RMS), pero con la evolución del negocio y la necesidad de integración con la facturación electrónica, se migró hacia un sistema desarrollado a la medida llamado OVST. Este sistema permite manejar operaciones específicas del giro estético, pero funciona de manera aislada respecto a otras unidades. Es aquí que resulta esencial poner bajo la lupa el flujo de negocio del inventario, es decir, el camino que recorre cada producto desde que ingresa al grupo (Ingreso o transferencia), pasa por las bodegas y salones, se consume en un servicio o se vende al cliente final, y finalmente se refleja en los reportes de gestión. Al entender con detalle este recorrido podremos identificar los puntos donde la información se fragmenta, se duplica o se pierde, y, a partir de ello, establecer las mejoras necesarias para lograr una visión unificada, oportuna y confiable de las existencias en todas las sociedades de Grupo Probe.

Esta situación genera las siguientes problemáticas:

- **Imprecisión en los inventarios:** Es común que los saldos de inventario no reflejen con precisión el consumo real de productos, especialmente en Vidals y Sento, donde los insumos (como esmaltes, tintes, cremas, etc.) son utilizados directamente en la prestación de servicios y no siempre se registran adecuadamente como consumo interno.
- **Duplicación de información y errores de conciliación:** Las transacciones deben registrarse en sistemas diferentes, lo que incrementa el riesgo de errores humanos, retrasos en la actualización de saldos y diferencias entre los sistemas.
- **Limitación en el análisis y toma de decisiones:** La falta de un sistema consolidado impide generar reportes confiables sobre rotación de productos, consumo por servicio, pérdidas o necesidades de reposición, afectando directamente la capacidad de respuesta logística y el control financiero.

En este contexto, el reto de Grupo Probe es contar con una visión unificada y confiable del inventario, que abarque tanto los productos destinados a la venta como aquellos utilizados en servicios, y que permita una alimentación de datos automatizada, trazable y verificable. El inventario debe cumplir una doble función: como recurso productivo (consumo directo en servicios) y como producto comercial (venta directa en TBS). Esta problemática, además de obstaculizar la eficiencia operativa, tiene repercusiones económicas cuantificables: pérdidas por productos no registrados, Ingresos innecesarias, ruptura de stock, sobrecostos logísticos y decisiones basadas en datos incompletos o imprecisos. Por ello, el foco de esta investigación es el diseño e implementación de una solución de ingeniería de datos, orientada a integrar las fuentes de información existentes, depurar y unificar los registros, y establecer una base que permita el análisis avanzado del inventario en todo el grupo.

C. Planteamiento del problema



V. Hipótesis

A. Hipótesis nula

No existe una diferencia significativa, ni en el tiempo promedio de respuesta de las consultas analíticas de inventarios ni en el tiempo de carga de datos nuevos, cuando los cubos OLAP se generan a partir de vistas SQL que recorren la totalidad de los registros, en comparación con los cubos generados sobre un Data Warehouse en esquema de estrella.

B. Hipótesis alterna

La generación de cubos OLAP sobre un Data Warehouse en esquema de estrella alimentado con procesos ETL que permiten cargas incrementales reduce en al menos un porcentaje considerable de respuesta de las consultas analíticas de inventarios y disminuye el tiempo de carga de nuevos datos en comparación con la generación de cubos OLAP utilizando vistas SQL que requieren recargar todos los registros.

VI. Justificación

Justificación del Proyecto: Diseño de un Modelo Multidimensional y Construcción de Data Warehouse para Análisis de Inventarios del Grupo PROBE

Analizaremos cuáles serán los beneficios cualitativos y cuantitativos que genera la Construcción de data warehouse para análisis de inventarios para las sociedades de grupo probe impactando directamente la eficiencia operativa, la rentabilidad y la toma de decisiones estratégicas de la empresa.

Beneficios Cuantitativos

Los beneficios cuantitativos de este proyecto se traducirán en mejoras medibles y tangibles en las operaciones de inventario del Grupo PROBE:

- **Reducción del Tiempo de Disponibilidad de la Información:** Significativa disminución en el tiempo requerido para acceder a la información de inventario y generar reportes en tiempo real, lo que agilizará la toma de decisiones estratégicas y operativas.
- **Control Integral del Inventario:** Mejora sustancial en el control de inventario tanto para productos destinados a la venta como para aquellos utilizados en la prestación de servicios, garantizando la precisión de las existencias.
- **Integración y Estandarización de Unidades de Negocio:** Facilitar la integración de todas las unidades de negocio al replicar un modelo unificado, promoviendo la consistencia operativa y el flujo de información inter-sociedades.
- **Consolidación Unificada de la Información:** Obtener una visión consolidada y única de la información de inventario de todas las sociedades que conforman Grupo PROBE, eliminando silos de datos y ofreciendo una perspectiva integral.
- **Prevención de Pérdidas por Productos No Registrados:** Minimizar las pérdidas económicas asociadas a productos que no son debidamente registrados en el sistema, asegurando la trazabilidad completa del inventario.
- **Optimización de Ingresos y Reducción de Costos:** Ejercer un control preciso sobre los Ingresos, evitando adquisiciones innecesarias y optimizando la logística del negocio, lo que resultará en una reducción de costos operativos y una mejor gestión del capital de trabajo.

Beneficios Cualitativos

Los beneficios cualitativos, aunque no siempre monetizables directamente, son cruciales para la competitividad y el crecimiento sostenible del Grupo PROBE:

Mejora en la Agilidad y Flexibilidad Operacional: Al tener una visión clara y en tiempo real del inventario, Grupo Probe podrá responder de manera más rápida y eficiente a las demandas del mercado, a las fluctuaciones de stock y a las oportunidades de negocio, mejorando la agilidad en

la gestión de toda su cadena de suministro.

Toma de Decisiones Estratégicas y Oportunas: La disponibilidad de información consolidada y precisa permitirá a la gerencia tomar decisiones más informadas y oportunas, no solo en la gestión diaria del inventario, sino también en la planificación estratégica de Ingresos, promociones y lanzamiento de productos, así como en la expansión de servicios.

Mayor Confianza y Fiabilidad en los Datos: La consolidación y el control mejorado de la información de inventario aumentarán la confianza en la exactitud de los datos. Esto se traducirá en una mayor credibilidad de los reportes y análisis internos, facilitando una cultura de decisión basada en hechos.

Estandarización y Armonización de Procesos: La integración de las unidades de negocio y la replicación del modelo de datos fomentarán la estandarización de los procesos de inventario a lo largo de todas las sociedades. Esto reducirá las inconsistencias operativas y mejorará la eficiencia general.

Fortalecimiento de la Colaboración Inter-sociedades: Al contar con una plataforma unificada de información, se promoverá una mayor colaboración y comunicación entre las distintas sociedades de Grupo Probe (Vidals, Sento Luxury Spa & Salón, TBS Beauty Supply, Disprobe / Intrade), facilitando sinergias y una gestión más cohesionada.

Optimización del Capital de Trabajo : Aunque el control de Ingresos innecesarios tiene un impacto cuantitativo, cualitativamente, una mejor gestión del inventario permite liberar capital que de otra manera estaría inmovilizado en stock excesivo, lo que puede reinvertir en otras áreas estratégicas del negocio.

Mejora en la Satisfacción del Cliente (Interno y Externo):

- Interna: Los equipos de ventas y servicio tendrán acceso rápido a la disponibilidad de productos, mejorando su capacidad de atención y respuesta.
- Externa: Una gestión de inventario eficiente reduce las rupturas de stock y garantiza la disponibilidad de los productos, lo que se traduce en una mejor experiencia y satisfacción para los clientes finales de Grupo Probe.

Escalabilidad Sostenible: La mejora en la gestión de inventarios y la información consolidada proporcionarán una base sólida y escalable para el futuro crecimiento de Grupo Probé, permitiendo la integración de nuevas operaciones o líneas de negocio con mayor facilidad.

VII. Metodología de trabajo

1. Metodología Asociada al Proceso de Investigación

a. Comprensión y Levantamiento de Requerimientos

- **Reuniones y Entrevistas con Stakeholders:** utilizaremos la metodología scrum donde se realizarán sesiones de trabajo con los ingenieros a cargo de Grupo Probe (Vidals, Sento Luxury Spa & Salón, TBS Beauty Supply, Disprobe / Intrade) ,El objetivo es entender sus procesos actuales de inventario, sus desafíos, los reportes que actualmente utilizan y sus necesidades de información y análisis.
- **Revisión de Documentación Existente:** Analizaremos cualquier documentación relevante sobre los procesos de inventario actuales, flujos de trabajo, políticas y procedimientos de gestión de stock.
- **Análisis del Modelo de Datos Actual:** Revisaremos la arquitectura de la solución de inteligencia de negocios existente, incluyendo los cubos OLAP ,en Microsoft SQL Server. sus fuentes de datos y la lógica de negocio implementada.

b. Análisis de Datos (Diagnóstico)

- **Exploración de Datos del Cubo OLAP:** Utilizar las herramientas de análisis de Microsoft SQL Server (SSAS) y los cubos OLAP existentes para explorar los datos de inventario. Esto incluirá:
 - Verificación de la calidad y consistencia de los datos.
 - Identificación de patrones, tendencias y anomalías en los datos de inventario (ej., stock obsoleto, rotación lenta, discrepancias).
- **Validación de Indicadores y Reportes:** Evaluar la pertinencia y precisión de los indicadores de gestión de inventario y los reportes actuales. Se buscará determinar si estos procesos permiten una toma de decisiones efectiva en tiempo real.
- **Mapeo de Procesos Actual vs. Deseado:** Comparar los flujos de trabajo de inventario observados con los ideales, identificando los puntos de fricción y las oportunidades de automatización o mejora.

c. Definición de Oportunidades de Mejora y Soluciones (Diseño)

- **Propuesta de Optimización de Procesos:** Basado en el análisis, se formularán recomendaciones específicas para la optimización de los procesos de gestión de inventario, incluyendo posibles ajustes en la forma en que se registran, mueven o gestionan los productos.
- **Diseño de Nuevas Métricas:** En colaboración con los stakeholders, se definirán nuevos indicadores clave de rendimiento y métricas que capturen de manera más efectiva el estado y la eficiencia del inventario en tiempo real.
- **Propuesta de Mejoras en el Modelo Dimensional:** Identificar si se requieren ajustes o extensiones al modelo dimensional actual para soportar los nuevos requerimientos de análisis definidos. Esto podría incluir la adición de nuevas dimensiones o atributos, o la redefinición de hechos.

- **Plan de Implementación (ETL y Datawarehouse):** Desarrollar un plan conceptual para implementar las mejoras identificadas en el proceso ETL (Extracción, Transformación y Carga) , asegurando que la información se consolide de manera eficiente y se actualice en tiempo real.

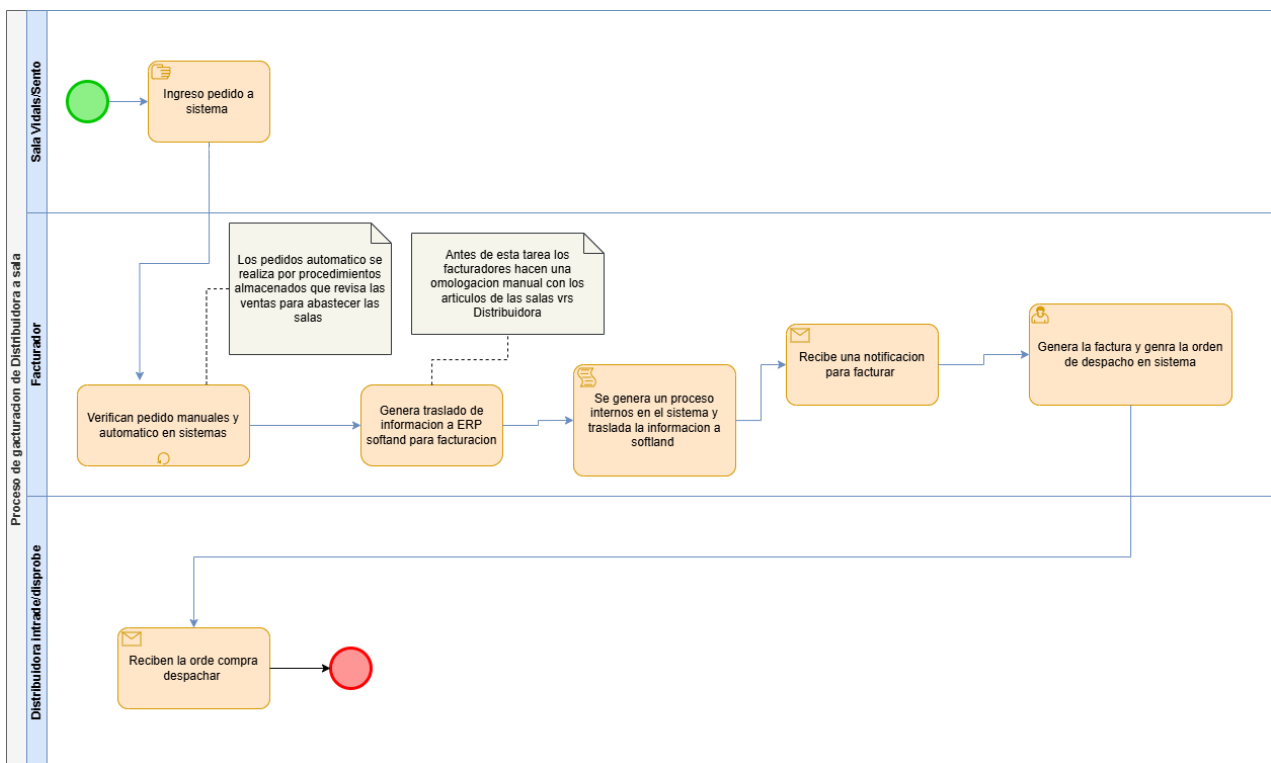
d. Validación y Planificación Futura

- **Sesiones de Validación:** Presentar los hallazgos, las propuestas de mejora y los diseños conceptuales a los stakeholders en nuestro caso con los ingenieros a cargo clave para obtener su retroalimentación y validación.
- **Roadmap de Implementación:** Definir un plan de acción con hitos claros, responsabilidades y una estimación de recursos y tiempos para la implementación de las soluciones propuestas.
- **Consideraciones de Escalabilidad:** Asegurar que las soluciones propuestas no solo resuelvan los desafíos actuales, sino que también sean escalables para soportar el futuro crecimiento y la expansión de Grupo Probe.

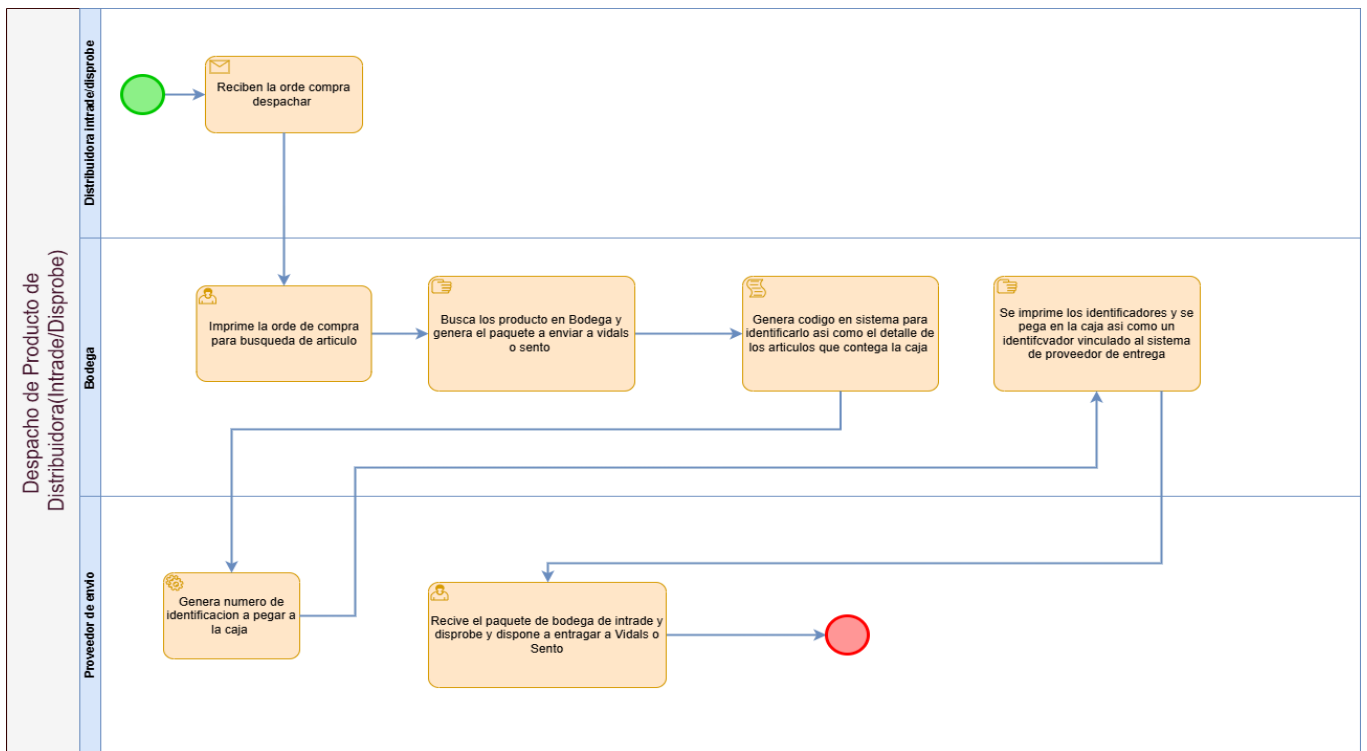
Esta metodología permitirá abordar de manera integral el desafío del inventario en Grupo Probe, transformando la data en información valiosa y el activo más importante: el conocimiento para la toma de decisiones.

Proceso de negocio

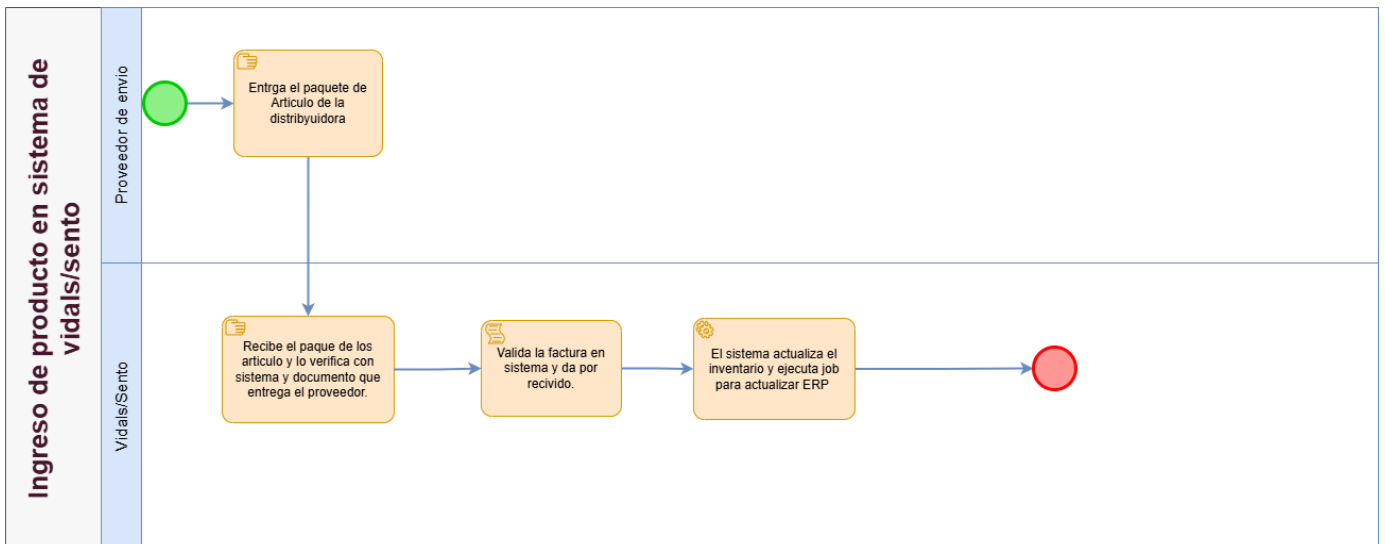
- BPMN Facturación de Distribuidora a sala Vidals/Sento



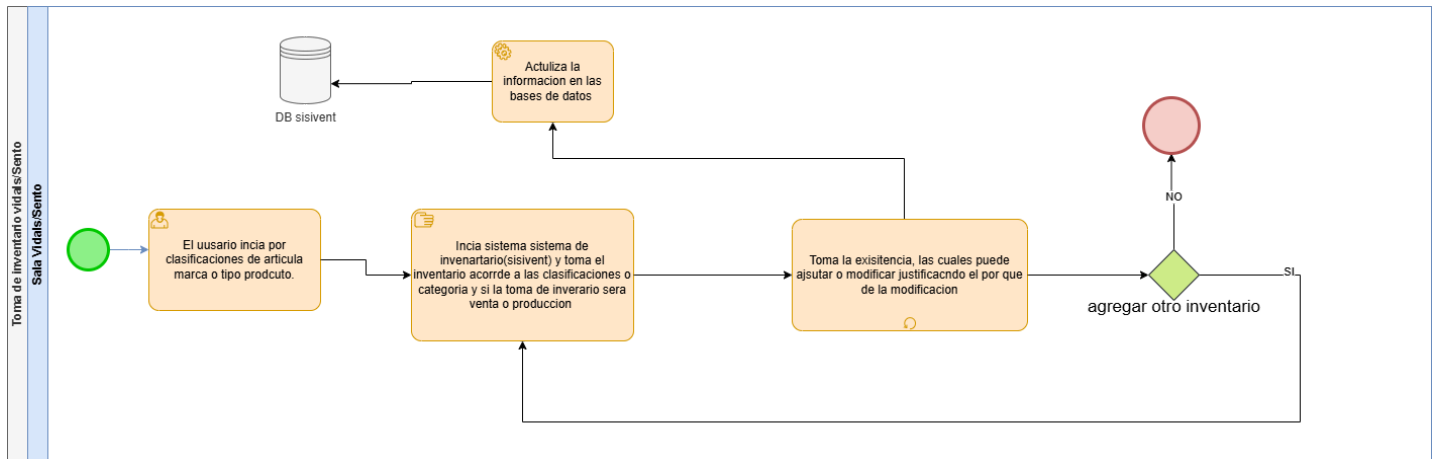
- BPMN Despacho de Producto de Distribuidora(Intrade/Disprobe)



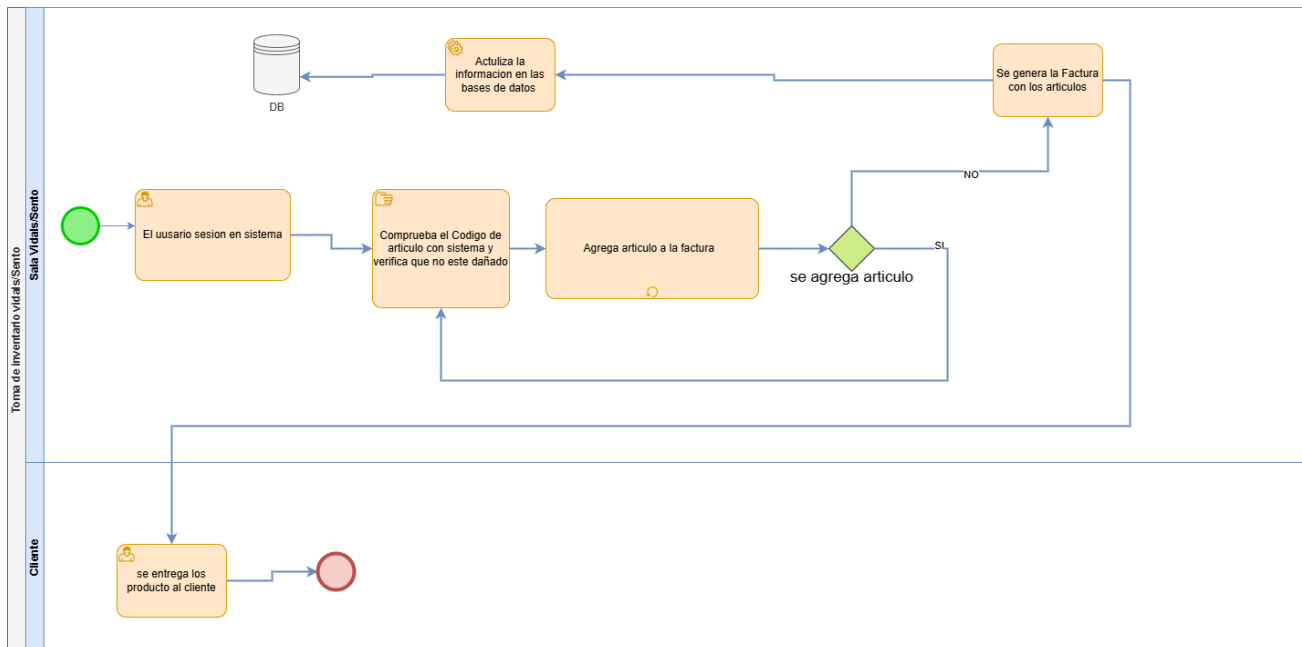
- BPMN Ingreso de Producto en sistema de Vidals o Sento



- BPMN Control de inventario(para venta/ para uso de servicio)



- BPMN venta de producto



VIII. Descripción de la propuesta de solución

Grupo Probe enfrenta actualmente una serie de dificultades en la forma en que gestiona sus inventarios, debido a que cada una de sus sociedades opera con sistemas diferentes que realizan ajustes antes de unificar la información en Softland.

Esto ha provocado errores en los registros, duplicación de datos, retrasos en la información y problemas para tomar decisiones basadas en cifras confiables. Además, el sistema que se usa actualmente para analizar la información se ha quedado corto frente al crecimiento del negocio, pues depende de procesos lentos, poco flexibles y que consumen muchos recursos.

La solución que se propone consiste en construir una plataforma central que reúna toda la información de inventarios en un solo lugar siendo esto nuestro enfoque de los procesos de negocios que tiene Grupo Probe. Esta plataforma será como una “fuente única de verdad”, donde se concentrará y organizará todo lo que ocurre en los salones Vidals y Sento, en las tiendas TBS, y en las operaciones de Intrade e Disprobe siendo esta última sociedad punto de enfoque. La información será actualizada de forma automática, sin necesidad de hacerlo todo desde cero cada vez, lo que permitirá tener datos frescos cada día, sin depender de procesos manuales o lentos.

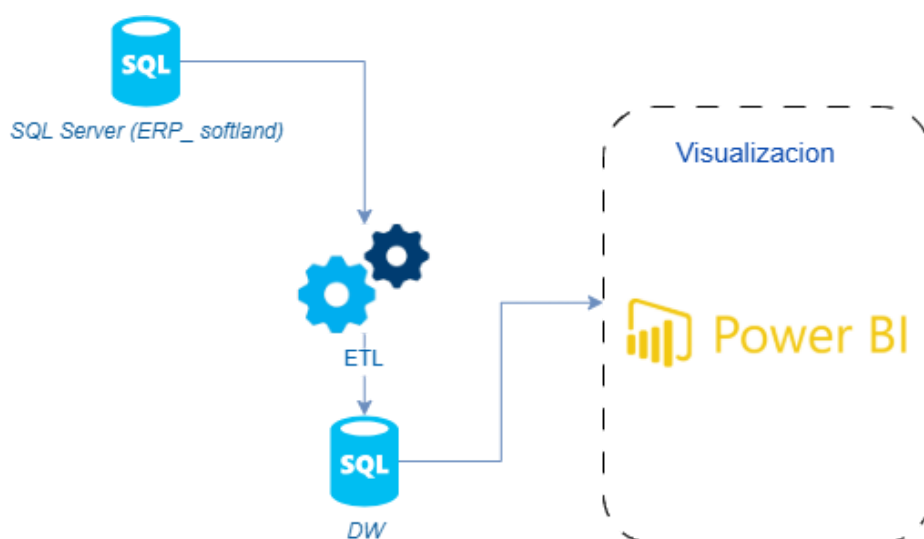
Para lograr esto, se conectarán al sistema (Softland ERP) a esta plataforma central. La información será ordenada, corregida y adaptada para que todos hablen el mismo idioma: mismo nombre para los productos, misma unidad de medida, mismas reglas para todos. Así, aunque los datos vengan de distintos lugares, podrán compararse y analizarse sin confusión.

Con esta solución, se podrá saber exactamente cuántos productos hay disponibles, cuánto se ha vendido, cuánto se ha usado en servicios, y cuándo se necesita reponer stock. Además, permitirá identificar pérdidas, hacer seguimiento al inventario entregado como cortesía o promoción, y conocer la rentabilidad real de cada artículo o servicio. Toda esta información podrá ser elaborada y disponible en reportes claros y visuales, listos para usarse por las áreas de Ingresos, finanzas, operaciones o dirección general.

Los beneficios que esta propuesta traerá son tangibles. Por ejemplo, los tiempos de procesamiento bajarán de varias horas a menos de una hora por día. Se reducirá la pérdida de productos por falta de control, lo que puede representar ahorros importantes (entre un 5% y 10% de reducción de pérdidas). Se evitarán ingresos innecesarios gracias a una mejor planificación, y se facilitará el trabajo de las áreas de datos y sistemas, ya que no tendrán que intervenir constantemente para resolver errores o generar reportes especiales.

Desde el punto de vista organizacional, esta solución también representa una modernización importante. Disminuye la dependencia de vistas SQL específicas por sociedad y permite que los usuarios operativos y gerenciales puedan autogestionar sus consultas, sin requerir asistencia técnica constante. Finalmente, esta arquitectura es escalable, lo que significa que se podrán integrar nuevas sociedades, indicadores o sistemas sin necesidad de rediseñar el modelo desde cero.

Este diagrama representa a grandes rasgo la solución propuesta para grupo probe :

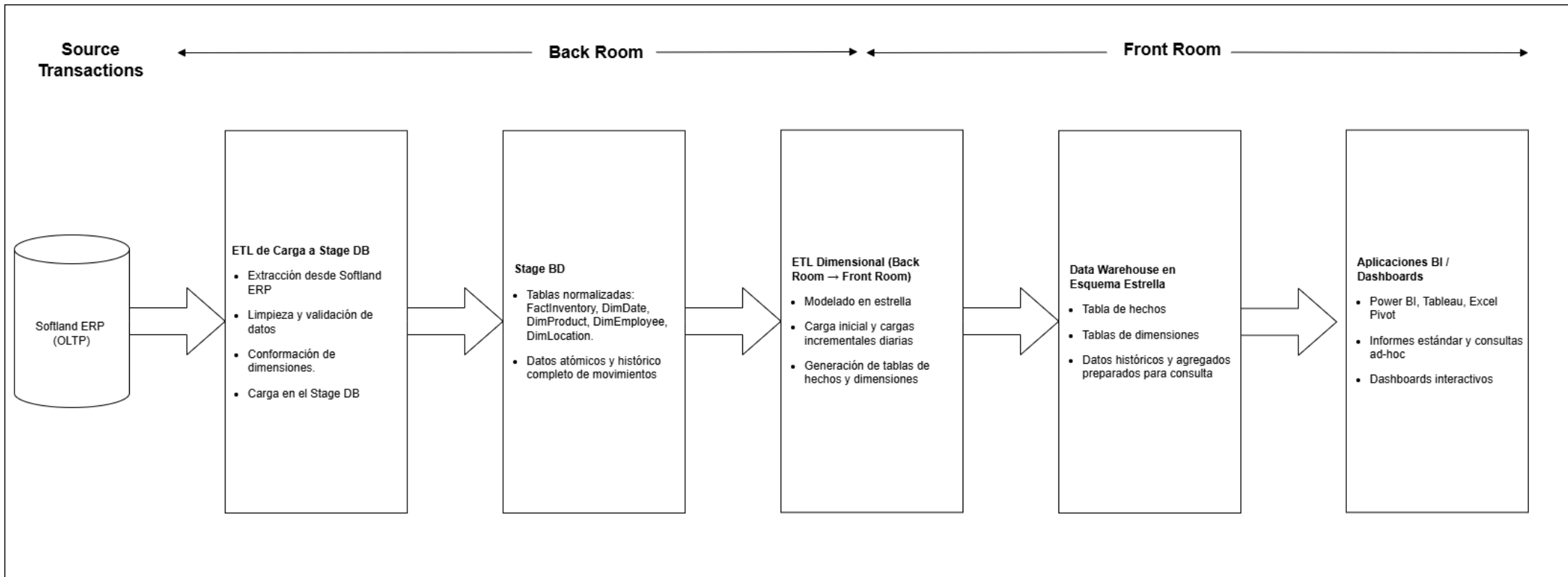


IX. Tecnologías a utilizar

Herramienta	Versión	Requisitos	Justificación
SQL Server	2022 Developer	HDD: 6GB RAM 4GB Procesador: 2Ghz o superior S.O: Windows 10 o superior	Se utiliza la version developer debido a que es una versión gratuita por motivos de desarrollo por parte del equipo y es el principal motor de base de datos de la empresa Probe.
DBeaver	Community 25.1.0	S.O: Windows 8 o superior	Sistema gestor de base de datos que integra compatibilidad con diferentes motores de base de datos, proporciona una interfaz para visualizar diagramas relacionales de acuerdo a las tablas creadas. Es una herramienta gratuita, la cual permite utilizar conexiones de múltiples motores de base de datos.
Python	3.13.4		Lenguaje seleccionado para desarrollar los ETL del Data Warehouse, es un lenguaje con una curva de aprendizaje fácil y permite un mantenimiento sostenible con pocas horas de práctica del lenguaje.

Visual Studio Code			Principal IDE para desarrollar
Apache Spark	4.0.0	Python (Opcional) Docker	Herramienta utilizada para crear los ETL que se conectarán al DW y recolectarán la información de los sistemas transaccionales para este.
Power BI Desktop	2.143.1378.0 64-bit	S.O: Windows 10 HDD: 2.9GB RAM: 2 GB Procesador: 1Ghz o superior	Herramienta para presentar y comprobar el trabajo realizado en el proceso de generación del Data Warehouse. Power BI se utiliza por la empresa Probe para visualizar tableros que ya poseen.
Git	web	Conexión a internet	Herramienta utilizada para realizar el control de versiones del proyecto, almacena los archivos del proyecto.
Google Docs	web	Conexión a internet	Herramienta para redactar la documentación del proyecto e implementar la descripción del modelo físico de la base de datos.
Draw.io	Web	Conexión a internet	Herramienta gratuita para realizar diagramas relevantes de bases de datos. Sean diagramas físicos o de entidad relación.
Trello	web	Conexión a internet	Herramienta para coordinar y ordenar el trabajo del equipo. Mediante esta herramienta se gestiona el proyecto. Es fácil de utilizar y representa una forma de visualizar los avances del proyecto.
Google Meet	web	Conexión a internet	Principal vía de comunicación entre el equipo de especialización y la empresa Probe. En esta plataforma se realizan reuniones de seguimiento del proyecto por parte del equipo

X. Diagrama arquitectónico de la solución



XI. Descripción de cada componente de la solución

La arquitectura propuesta sigue un flujo típico de un sistema de Business Intelligence (BI) y Data Warehousing, desde las transacciones de origen hasta las aplicaciones de consumo de datos. Se divide conceptualmente en "Source Transactions", "Back Room" y "Front Room".

Descripción de Componentes de la Solución de Data Warehousing según modelo arquitectónico

1. Source Transactions: Softland ERP (OLTP)

Sistema transaccional de tipo OLTP (Online Transaction Processing) de Softland. Es la fuente primaria de datos operativos que utilizaremos para generar la base para el datawarehouse.

- Contendrá los datos detallados de las operaciones diarias de la organización, como ventas, inventarios, transacciones de empleados, etc. Está optimizado para la inserción, actualización y eliminación rápida de registros individuales.
- Este es el sistema de origen existente. La estrategia será extraer datos de él sin modificar su funcionalidad principal.

2. Back Room: ETL de Carga a Stage DB

Proceso de ETL (Extract, Transform, Load) encargado de la extracción inicial y la preparación de los datos desde el ERP hacia la base de datos de staging.

1. **Extracción desde Softland ERP:** Conecta con el sistema Softland ERP para obtener los datos brutos.
2. **Limpieza y validación de datos:** Aplica reglas para identificar y corregir errores, inconsistencias o datos faltantes (ej., formato de fechas, valores nulos, duplicados).
3. **Conformación de dimensiones:** Prepara los datos para su uso en dimensiones futuras, asegurando la consistencia de los atributos maestros.
4. **Carga en el Stage DB:** Almacena los datos limpios y validados temporalmente en la base de datos de staging.

3. Back Room: Stage BD (Base de Datos de Staging)

Base de datos temporal diseñada para almacenar los datos extraídos y pre-procesados antes de ser cargados en el Data Warehouse.

Tablas normalizadas: Almacena los datos en un formato que, aunque puede ser ligeramente desnormalizado para eficiencia, mantiene la estructura cercana a la fuente, facilitando la limpieza y transformación..

Datos atómicos y el histórico completo de movimientos: Contiene el nivel de detalle más granular de los datos y, potencialmente, un historial completo de cambios si se requiere para la trazabilidad.

4. Back Room: ETL Dimensional (Back Room -> Front Room)

Proceso de ETL avanzado que transformará los datos del Stage DB en un modelo dimensional (esquema estrella o copo de nieve) y los carga en el Data Warehouse.

- ◆ **Modelado en estrella:** Aplica las transformaciones necesarias para conformar los datos en hechos y dimensiones según el diseño del esquema estrella.
- ◆ **Carga inicial y cargas incrementales diarias:** Ejecuta una carga completa inicial y luego procesos incrementales periódicos (diarios) para mantener el Data Warehouse actualizado con los nuevos datos.
- ◆ **Generación de tablas de hechos y dimensiones:** Construye las tablas de hechos (que contienen las métricas numéricas y claves foráneas a las dimensiones) y las tablas de dimensiones (que contienen los atributos descriptivos).

5. Front Room: Data Warehouse en Esquema Estrella

Base de datos relacional optimizada para consultas analíticas, implementada siguiendo un esquema estrella (o copo de nieve) para facilitar la navegación y el rendimiento de las consultas BI.

Tablas de hechos: Contienen las métricas cuantitativas clave del negocio (ej., cantidad vendida, monto de ventas) y claves foráneas a las tablas de dimensiones.

Tablas de dimensiones: Proporcionan el contexto para los hechos, incluyendo atributos descriptivos sobre entidades clave (ej., clientes, productos, tiempo, ubicación, empleados).

Datos históricos y agregados preparados para consulta: Almacena datos a lo largo del tiempo y puede contener resúmenes o agregaciones pre-calculadas para acelerar las consultas de informes y dashboards.

6. Front Room: Aplicaciones BI / Dashboards

Conjunto de herramientas de Business Intelligence y visualización de datos utilizadas por los usuarios finales para explorar, analizar y presentar la información del Data Warehouse.

- **Power BI, Tableau, Excel Pivot:** Ejemplos de herramientas que permiten a los usuarios crear informes interactivos, paneles de control y realizar análisis ad-hoc.
- **Informes estándar y consultas ad-hoc:** Capacidad para generar reportes predefinidos y permitir a los usuarios construir sus propias consultas para explorar los datos.

- **Dashboards interactivos:** Proporcionan una vista consolidada y visual de los indicadores clave de rendimiento (KPIs) y tendencias, permitiendo la exploración de datos a diferentes niveles.


XII. Cronograma de actividades

Las actividades principales para el desarrollo del proyecto son las siguientes:

Fecha de realización	Actividad Principal	Actividades Secundarias
9-05-2025 al 11-06-2025	Presentación y definición de requerimientos	<ul style="list-style-type: none"> ● Presentación del equipo del proyecto ● Definición del alcance y los objetivos del proyecto ● Recopilación de requerimientos del negocio ● Evaluación de posibles fuentes de datos ● Identificación de la tecnología utilizada ● Elaboración de planificación de proyecto ● Definición de las tareas, responsabilidades mediante trello ● Planificación de reuniones para presentación de avances ● Levantamiento del entorno de trabajo ● Presentación de avances y feedback de usuarios
16-06-2025 al 15-07-2025	Diseño del Data Warehouse	<ul style="list-style-type: none"> ● Identificación de procesos de negocio por prioridad de negocio ● Definición de la granularidad de los procesos ● Identificación de las fuentes de datos necesarias ● Diseño de dimensiones del data warehouse ● Diseño de de tablas de hechos del data warehouse ● Creación de documentación del modelo físico del DW ● Generación de documentación. ● Presentación de avances y feedback de usuarios
16-07-2025 al 25-07-2025	Construcción del Data Warehouse	<ul style="list-style-type: none"> ● Creación de tablas de dimensiones ● Creación de tablas de hechos ● Generación de roles y usuarios ● Presentación de avances y feedback de usuarios
11-08-2025 al	Construcción de	<ul style="list-style-type: none"> ● Codificación de la extracción

17-10-2025	herramienta ETL	<ul style="list-style-type: none"> • Desarrollo de la transformación de datos • Limpieza de datos • Diseño y construcción de carga • Pruebas al ETL • Presentación de avances y feedback de usuarios
20-10-2025 al 03-11-2025	Generación de tablero	<ul style="list-style-type: none"> • Creación de un tablero de prueba de la información relevante • Pruebas de integración de componentes de la solución • Redacción de documentación de la solución • Presentación de la solución y feedback de usuarios • Entrega de la herramienta a la empresa

A continuación se desarrolla un diagrama de Gantt, donde se puede visualizar la planificación general del proyecto, tomando en cuenta un flujo de trabajo exitoso. Lo que significa que los tiempos pueden variar de acuerdo a ciertas condiciones siendo esta una proyección del trabajo a realizar.

 Diagrama de Gantt


XIII. Presupuesto

Proyecto: Construcción de data warehouse para análisis de inventarios para las sociedades de grupo probe							
ITEM	DESCRIPCION	HORAS	UNIDAD	PRECIO UNITARIO (\$)	SUB TOTAL (\$)	TOTAL (\$)	REFERENCIA
1	RECURSOS HUMANOS					\$ 5,133.33	
1.1	Business Analyst (Levantamiento de requisitos, modelado lógico)	176	1	\$ 6.25	\$ 1,100.00		Link
1.2	Data Engineer(Desarrollo de ETL y Data Warehouse)	176	2	\$ 8.33	\$ 2,933.33		Link
1.4	QA Developer (Pruebas de integridad, rendimiento y validación)	176	1	\$ 6.25	\$ 1,100.00		Link
2	SOFTWARE / HARDWARE	CANTIDAD				\$ 10,953.60	
2.1	Servidor on-premise (Equipo rackmount, Dell PowerEdge R440)	1	UND	\$ 5,134.95	\$ 5,134.95		Link
2.2	Almacenamiento SAN (5 discos Seagate SkyHawk 4 TB + chasis SAN)	5	UND	\$ 139.53	\$ 697.65		Link
2.3	Windows Server Standard (16-core, licencia producción)	1	Licencia	\$ 1,176.00	\$ 1,176.00		Link
2.4	SQL Server Standard (2-core pack para BI)	1	Licencia	\$ 3,945.00	\$ 3,945.00		Link
3	CONSTRUCCION	HORAS				\$ 1,166.20	
3.1	Instalación (Montaje de hardware y software en intranet)	40	1	\$ 8.33	\$ 333.20		Link
3.2	Configuración DWH(Parametrización de SSIS, SSAS y base relacional)	60	1	\$ 8.33	\$ 499.80		Link
3.3	Capacitación(Transferencia de conocimiento al equipo interna)	20	2	\$ 8.33	\$ 333.20		Link
	SUB-TOTAL					\$ 17,253.13	
5.8	Fondo de contingencia (10% del proyecto)	1	Fondo	\$ 1,725.31	\$ 1,725.31	\$ 1,725.31	
	TOTAL					\$ 18,978.45	

XIV. Anexos

<https://www.grupoprobe.com/>

Diagrama de Gantt

 Diagrama de Gantt

Trello del equipo

<https://trello.com/b/4JlPo6Pf/especializacion-de-bd>

Presupuesto

[presupuesto_data_warehouse.xlsx](#)