



Westfield
Business
School



Maestría en

ADMINISTRACIÓN DE EMPRESAS

**Tesis previa a la obtención de título de Magister en
Administración de Empresas**

AUTORES:

Abg. Sara Paez Vallejo
Ing. Nathaly Michelle Coello Villagran
Lcda. Karol Alejandra Mise Vasco
Adm. David Hernan Bermeo Rocha
Ing. Richard Francisco Castillo Gallardo

TUTOR: Ing. José Luis Perez Galán

Propuesta de creación de “DataLab BI: Plataforma que centraliza toda la información de Visionlab Ecuador en un único portal, acelerando los procesos administrativos y potenciando la eficiencia de la fuerza de ventas.”

PROYECTO DE TITULACIÓN – FIN DE MÁSTER
MAESTRÍA EN ADMINISTRACIÓN DE EMPRESA ENL

DataLab BI: Plataforma que centraliza toda la información de Visionlab Ecuador.

Por

Abg. Sara Paez Vallejo

Ing. Nathaly Michelle Coello Villagran

Lcda. Karol Alejandra Mise Vasco

Adm. David Hernan Bermeo Rocha

Ing. Richard Francisco Castillo Gallardo

Octubre 2025


Aprobado



Cristian Melo

Presidente(a) del Tribunal Universidad Internacional del Ecuador

Yo, Cristian Javier Melo González e Ignacio Maroto, declaramos que, personalmente conocemos que los graduandos Abg. Sara Paez Vallejo, Ing. Nathaly Michelle Coello Villagran, Lcda. Karol Alejandra Mise Vasco, Adm. David Hernan Bermeo Rocha, Ing. Richard Francisco Castillo Gallardo, son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.

A handwritten signature in blue ink, appearing to read 'CRISTIAN', with a horizontal line drawn underneath it.

Cristian Melo

Coordinador MBA UIDE

A handwritten signature in blue ink, appearing to read 'Ignacio', with a horizontal line drawn underneath it.

Ignacio Maroto

Provost WBS

Autoría del Trabajo de Titulación

Nosotros. Abg. Sara Paez Vallejo, Ing. Nathaly Michelle Coello Villagran, Lcda. Karol Alejandra Mise Vasco, Adm. David Hernan Bermeo Rocha, e Ing. Richard Francisco Castillo Gallardo, , declaro bajo juramento que el trabajo de titulación titulado **DataLab BI: Plataforma que centraliza toda la información de Visionlab Ecuador.** es de nuestra autoría y exclusiva responsabilidad legal y académica; que no ha sido presentado anteriormente para ningún grado o calificación profesional, habiéndose citado las fuentes correspondientes y respetando las disposiciones legales que protegen los derechos de autor vigentes.



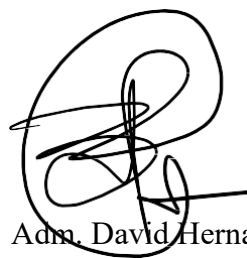
Abg. Sara Paez Vallejo
Correo electrónico:
spaezsocialesd@gmail.com



Ing. Nathaly Michelle Coello Villagran,
Correo electrónico: Naty.c_16@hotmail.com



Lcda. Karol Alejandra Mise Vasco
Correo electrónico: karolmise15@gmail.com



Adm. David Hernan Bermeo Rocha
Correo electrónico: davher2009@hotmail.com



Ing. Richard Francisco Castillo Gallardo Correo
electrónico: richard.cas.gall@gmail.com

Autorización de Derechos de Propiedad Intelectual

Nosotros, Abg. Sara Paez Vallejo, Ing. Nathaly Michelle Coello Villagran, Lcda. Karol Alejandra Mise Vasco, Adm. David Hernan Bermeo Rocha, e Ing. Richard Francisco Castillo Gallard, en calidad de autores del trabajo de investigación titulado **DataLab BI: Plataforma que centraliza toda la información de Visionlab Ecuador**, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que me pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, Octubre 2025



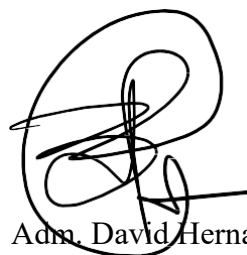
Abg. Sara Paez Vallejo
Correo electrónico:
spaezsocialesd@gmail.com



Ing. Nathaly Michelle Coello Villagran,
Correo electrónico: Naty.c_16@hotmail.com



Lcda. Karol Alejandra Mise Vasco
Correo electrónico: karolmise15@gmail.com



Adm. David Hernan Bermeo Rocha
Correo electrónico: davher2009@hotmail.com



Ing. Richard Francisco Castillo Gallardo Correo
electrónico: richard.cas.gall@gmail.com

Dedicatorias y Agradecimientos

A nuestras familias, por ser el pilar que nos ha sostenido durante todo este proceso, por su apoyo incondicional, paciencia y amor que nos dieron la fuerza para alcanzar esta meta.

A nuestros compañeros de equipo, con quienes compartimos ideas, retos, aprendizajes y largas horas de trabajo; su compromiso, compañerismo y espíritu colaborativo hicieron posible que este proyecto tome forma y se convierta en una realidad.

Y a la experiencia vivida durante este programa de posgrado, que nos retó a salir de la zona de confort, a crecer profesional y personalmente, y a descubrir el valor del aprendizaje conjunto para transformar los desafíos en oportunidades.

Tabla de contenido

Presentación de la Empresa	6
Nombre de la Empresa	6
Misión y Visión de la Empresa	6
Giro de Negocio:	6
Sector y mercados	7
Cifras del negocio	7
Objetivos del Proyecto: Definición Inicial del Problema	8
Nombre del proyecto	8
Objetivo del proyecto	8
Objetivos específicos	9
Áreas y Personas que Demandan la Información y uso Actual	9
Áreas que demandan la información:	9
Uso que hacen actualmente	9
Áreas con la Información Actual y Necesidades Detectadas	10
Ventas:	10
Marketing:	11
Comex:	12
Metodología del proyecto	13
Alcance y límites	14
Mapa de actores y responsabilidades	15
Análisis de riesgos y controles básicos	15
Definición de los KPI's Relevantes	16
Ventas:	16
Marketing:	17
Comex:	18
Origen de la información y recurrencia de carga: interna y externa	19
Fuentes internas	19
Fuentes externas	19
Frecuencia de extracción	19
Conectores y técnicas	20
Sistemas y destino de la carga	20
1. Carga inicial completa	23
2. Carga incremental (CDC – Change Data Capture)	23
3. Micro-batch	24
4. Cargas en ventana (window loads)	24
Almacenamiento en data warehouse y datamarts	25
Modelado dimensional: esquema estrella y copo de nieve	25
Catálogo de tablas	25
Carencias y problemas en la carga de datos	27
Fuentes incompletas o inconsistentes: datos faltantes en CRM, duplicados en ERP.	27
Fallos de transformación: reglas de negocio no aplicadas, errores de tipado.	27

Retrasos y cuellos de botella: ETLs que tardan > 2 horas, bloqueos de conexiones.	27
Calidad de datos: sin perfiles de calidad (validaciones, limpieza, estandarización).	27
Calidad de datos: sin perfiles de calidad (validaciones, limpieza, estandarización).	27
Informes que se generan actualmente, necesidades detectadas y oportunidades	28
Inventario de informes actuales	28
KPIs en uso	28
Brechas de información detectadas	28
Oportunidades de mejora	29
Procesos de carga de datos (ETL, data lake)	29
Extracción: (CMDB, Scripts, CUCM, SolarWinds, archivos Excel de facturación)	29
Transformación: Limpieza de datos, validación, homologación de formatos.	30
Carga: Datawarehouse y visualización a través de dashboards (Power BI, SAS, etc.).	30
Frecuencia: Diaria o mensual, según fuente.	31
Herramientas: SAS Data Integration Studio, Sas Visual Analytics	31
Información y calidad de la información disponible	32
Auditorías Mensuales Cruzando datos entre fuentes	32
Validación por la dirección de capacidades analíticas	32
Documentación de campos: Tipo de Dato, origen y frecuencia	33
Uso de logs para trazabilidad.	33
Horarios de carga y accesibilidad	33
Variables disponibles y acceso al Data Warehouse	34
Dominio Comercial (Ventas)	34
Dominio Marketing	34
Dominio Cliente	35
Sistemas implicados en los procesos	36
Recursos de IT o externos empleados	37
Análisis de herramientas y selección según el ecosistema	38
• Inventario de herramientas BI disponibles (Power BI, SAS, etc.)	38
• Comparativa técnica: integración, visualización, escalabilidad	39
• Compatibilidad con arquitectura actual (CRM, ERP, Data Warehouse)	39
• Evaluación de licencias, costos y facilidad de uso	39
• Selección de la herramienta óptima (justificación técnica y estratégica)	40
Selección de usuarios y perfiles de uso por dashboard	41
Clasificación de perfiles de usuario	41
Usuarios por área y dashboards asociados	41
Dashboards creados	42
Definición de niveles de acceso y permisos	43
Necesidades de información y accesos por departamento y perfil	44
Relevamiento de necesidades por área (ventas, marketing, operaciones, dirección)	44
Variables críticas para cada departamento (KPIs específicos)	45
Herramientas de control de acceso y seguridad de datos	46
Revisión y cumplimiento de los objetivos iniciales	46
a. Objetivos iniciales definidos	46

b. Estrategia planteada hasta ahora	46
c. Comparación: objetivos vs. estrategia definida	47
d. Recomendaciones y próximos pasos	47
Identificar situación actual del departamento	48
Estado actual del uso de datos en Ventas, Marketing y Comex	48
Nivel de digitalización e integración de fuentes (CRM, ERP, Excel, etc.)	48
Capacidades actuales del equipo en análisis de datos	48
Herramientas actualmente utilizadas y sus limitaciones	49
Problemáticas derivadas de decisiones sin base analítica	49
Definición de las carencias en formación y recursos	49
Coordinación de los flujos entre otras áreas	51
• Flujo de datos entre Ventas, Marketing, Comex y TI	51
• Procesos de consolidación de información (CRM → BI)	52
• Puntos críticos donde hay fricción o pérdida de eficiencia	52
• Propuestas para sincronización de procesos y reportes	53
• Roles responsables en cada área para garantizar trazabilidad	53
Definir el organigrama, dependencias de otras áreas y solución de problemas	54
• Roles clave	55
Recursos externos destinados: subcontratación	56
Business Case de costes y beneficios	58
1. Costos de Implementación (One-Time Costs)	58
5. Retorno de Inversión (ROI) y Periodo de Recuperación	60
Parte II	64
Seleccionar e identificar fuentes de datos tanto externas como internas, estructuradas o semiestructuradas que puedan ser utilizadas	64
Importancia de la taxonomía de fuentes	64
1. Fuentes internas estructuradas	64
2. Fuentes internas semiestructuradas	64
3. Fuentes externas: estructuradas y semiestructuradas	65
Estrategias de ingestión, sincronización y gobernanza	65
Encontrar y definir la relación de los datos para ayudarnos a resolver y mejorar nuestra propuesta inicial.	66
Entender y justificar para nuestro proyecto el impacto que el big data puede aportar y en qué áreas, departamentos y procesos de negocio.	67
1. Ventas	67
2. Marketing	67
3. Comex	67
4. Dirección General	68
5. Procesos transversales	68
Analizar las alternativas de los diferentes proveedores propuestos.	68
1. Google Cloud (BigQuery)	68
2. Amazon Web Services (AWS)	69
3. Microsoft Azure	69
4. IBM Watson/Cloud	70

Comparación Estratégica	70
HACER UNA SELECCIÓN DE NUESTRA ALTERNATIVA MEJOR DE PROVEEDOR	70
DEFINICIÓN DE MODELOS DE CLOUD COMPUTING COMO SaaS, PaaS o IaaS	72
Modelo de Cloud Computing para DataLab BI	72
1. PaaS (Platform as a Service)	72
2. SaaS (Software as a Service)	72
3. IaaS (Infrastructure as a Service)	72
flujo de información de nuestra arquitectura big data con el data warehouse o estructura alternativa de BI alternativa	73
SEMANA 8	75
1. Selección de Frameworks	75
Análisis simplificado de los frameworks	75
Selección de Framework para Datalab BI según análisis	75
2. Otros elementos requeridos del ecosistema del big data: lenguajes de programación, base de datos, etc.	77
SEMANA 9	79
1. Identificación y definición del objetivo de aplicación de Data Science y Machine Learning en DataLab BI	79
2. Fuentes de Datos Utilizados en DataLab BI	80
a. Fuentes internas estructuradas	80
b. Fuentes externas estructuradas	80
c. Fuentes externas semiestructuradas	80
3. área de aplicación y departamentos de negocio implicados	81
4. mejora esperada del proyecto DATALAB BI	81
a. Centralización y automatización de la información	81
b. Eficiencia operativa y toma de decisiones ágil	82
c. Incremento en ingresos y rentabilidad	82
d. Mejor control y calidad de datos	82
e. Visibilidad integral y segmentación avanzada	82
f. Reducción de carga manual y errores	82
5. cuantificar la mejora en el área de aplicación a través de un business case	83
A. Punto de partida (baseline) y alcance	83
B. Supuestos cuantitativos (consistentes con el documento)	83
C) Beneficios tangibles por área (anuales)	83
D) Escenarios (sensibilidad)	84
E) Métricas de impacto por área (antes → después)	85
F) Finanzas del caso	85
G) Roadmap de captura del valor (hitos medibles)	86
H) Riesgos y mitigaciones (solo los críticos)	86
Fuentes de datos utilizadas.	87
Área de aplicación y departamentos de negocio implicados.	87
Identificación de la mejora esperada.	87
Cuantificar la mejora en el área de aplicación mediante un business case.	88
Bibliografía	88

Tabla de contenido

Ilustración 1	27
Ilustración 2	28
Ilustración 3	28
Ilustración 4	29
Ilustración 5	70
Ilustración 6	70
Ilustración 7	71

Presentación de la Empresa

NOMBRE DE LA EMPRESA

Visionlab S.A.

MISIÓN Y VISIÓN DE LA EMPRESA

Misión: Proporcionar productos innovadores que mejoren la calidad de vida de las personas ayudándoles a ver mejor.

Visión: Ser el líder más fiable en el cuidado ocular del Ecuador, reconocido por su innovación, calidad y compromiso con mejorar vidas.

GIRO DE NEGOCIO:

Visionlab se enfoca en el cuidado ocular, ofreciendo una variedad de productos y soluciones para mejorar la visión y la salud de los ojos. Su giro de negocio se puede resumir en dos áreas principales:

1. Cuidado Quirúrgico: Visionlab ofrece productos y equipos para procedimientos oftálmicos, como:
 - a) Implantables quirúrgicos: lentes intraoculares para cirugía de cataratas y corrección de presbicia.
 - i) Consumibles quirúrgicos: instrumentos y accesorios para procedimientos oftálmicos.
 - ii) Equipo quirúrgico: sistemas de facoemulsificación y herramientas de diagnóstico.
2. Cuidado de la Visión: Visionlab proporciona productos para corregir problemas de visión, como:

- a) Lentes de contacto: para hipermetropía, miopía, astigmatismo y presbicia.
- b) Soluciones desinfectantes: para limpiar y cuidar lentes de contacto.
- c) Productos para ojo seco: gotas lubricantes oculares para aliviar síntomas de ojo seco.

SECTOR Y MERCADOS

- Comercio mayorista de medicamentos, instrumentos médicos y materiales quirúrgicos.
- Clientes: farmacias independientes y cadenas, hospitales públicos y privados, clínicas oftalmológicas en todo el país.

Visionlab importa y distribuye estos productos en el país, trabajando con profesionales de la salud ocular y centros médicos para brindar soluciones innovadoras y de alta calidad para mejorar la visión y la salud ocular de los pacientes.

CIFRAS DEL NEGOCIO

Facturación: En el segmento de Cuidado de la visión, Visionlab ha mantenido un crecimiento sostenido durante los últimos tres años, alcanzando una facturación estimada de **USD 10,5 millones en 2024**. Este crecimiento refleja una mayor penetración en el canal hospitalario y óptico, así como una estrategia comercial basada en datos.

Número de empleados: La empresa cuenta actualmente con 38 colaboradores, distribuidos estratégicamente en Quito, Guayaquil y Cuenca, lo que permite una cobertura nacional eficiente.

Años de actividad: Visionlab. fue constituida el 1 de Febrero de 2003, acumulando más de 22 años de experiencia en el mercado ecuatoriano.

Presencia global: Como filial de Visionlab, la compañía forma parte de una red internacional con operaciones en *más de 7 países* en Sudamérica.

Indicadores financieros recientes (2023):

- **Crecimiento de ingresos netos:** +8,20 %
- **Crecimiento de activos totales:** +11,64 %
- **Margen neto:** +2,97 %

Estos indicadores reflejan una gestión financiera sólida y capacidad para invertir en proyectos de transformación digital.

PARTE I

OBJETIVOS DEL PROYECTO: DEFINICIÓN INICIAL DEL PROBLEMA

NOMBRE DEL PROYECTO

Propuesta de creación de “**DataLab BI**: Plataforma que centraliza toda la información comercial de Visionlab en un único portal, acelerando los procesos administrativos y potenciando la eficiencia de la fuerza de ventas.”

OBJETIVO DEL PROYECTO

Desarrollar la plataforma “DataLab BI” que consolide automáticamente todas las fuentes de datos de Visionlab y ofrezca un dashboard interactivo con indicadores clave de performance comercial, competencia y retorno de inversión.

OBJETIVOS ESPECÍFICOS

1. Incrementar los ingresos netos en un 5 a 10% durante el primer trimestre del año, identificando oportunidades de cross-sell y up-sell mediante el análisis de clientes de alto potencial en DataLab BI (Área de ventas).
2. Mejorar el retorno sobre la inversión (ROI) de las campañas en un 25 % durante el primer semestre, monitorizando en tiempo real métricas como costo por lead (CPL) y tasa de conversión (Área de Marketing).
3. Acortar en un 40 % el tiempo de generación de informes ejecutivos, mediante dashboards interactivos que eliminen tareas manuales y mejoren la calidad de datos (Área de COMEX).

ÁREAS Y PERSONAS QUE DEMANDAN LA INFORMACIÓN Y USO ACTUAL

ÁREAS QUE DEMANDAN LA INFORMACIÓN:

- Ventas
- Marketing
- Comex

USO QUE HACEN ACTUALMENTE

Ventas: Revisan diaria, semanal y mensualmente diversas fuentes de información para comprender sus áreas y enfocar los recursos en la mejora y el aumento de la cuota de mercado.

Gerente de Ventas: Realiza una revisión mensual del rendimiento a nivel nacional y por cada uno de los territorios de su equipo. Esto permite obtener una visión integral del equipo y su gestión en campo (Cardona, Morelos, & Caraballo, 2022).

Además, alerta sobre posibles riesgos u oportunidades en los territorios asignados a la fuerza de ventas para brindar feedback, acompañamiento, soporte y coaching, mejorando las habilidades del equipo.

Marketing: El Brand Manager usa información interna y externa para crear estrategias que aseguren el crecimiento de la cuota de mercado y fortalezcan la marca entre los médicos prescriptores (Cardona, Morelos, & Caraballo, 2022).

Comex: Consolidan información de fuentes internas y externas para controlar los KPI de la fuerza de ventas como ejecución, rendimiento, y venta directa. Esta información también se utiliza para calcular comisiones, premios y bonos.

ÁREAS CON LA INFORMACIÓN ACTUAL Y NECESIDADES DETECTADAS

VENTAS:

1. Crecimiento de ventas y participación de mercado

El objetivo central es alcanzar y superar los targets de ventas anuales y trimestrales con respecto al año anterior.

2. Ejecución impecable del plan comercial

Esto incluye la implementación de promociones, eventos y actividades tácticas como parte del Landing Plan. Se espera que cada representante ejecute con excelencia y reporte mensualmente sus acciones (Estrella, Segura, Martínez, Sánchez, & Bravo, 2023).

3. Cobertura y frecuencia de cuentas clave

Se monitorea la frecuencia mínima de visitas por segmento de cliente con un mínimo del 90% en diversos territorios del país (Cardona, Morelos, & Caraballo, 2022).

4. Productividad de la fuerza de ventas

Se establecen KPIs claros como:

- Visitas por día: mínimo 10 médicos + 4 farmacias por representante.
- Días efectivos de visita: mínimo 190 al año.
- Capacidad de ejecución: cumplimiento $\geq 95\%$ del plan de llamadas (Cardona, Morelos, & Caraballo, 2022).

5. Gestión de rebates y acuerdos comerciales

El equipo de ventas participa activamente en la negociación y seguimiento de programas de rebate por volumen de compra y sell-out (Estrella, Segura, Martínez, Sánchez, & Bravo, 2023).

6. Innovación y nuevos canales

Se promueve el desarrollo de nuevos canales como clínicas terapéuticas y subdistribuidores, así como el posicionamiento de productos del portafolio en cuentas no tradicionales

MARKETING:

1. Impulsar el crecimiento de ventas

El marketing actúa como un habilitador comercial clave, diseñando promociones, eventos y recursos para aumentar el volumen de ventas. Por ejemplo, se menciona el uso de recursos limitados para generar acciones que aseguren que el paciente “siempre compre más”. También se promueve la ejecución de eventos como el Visionlab Contact Lenses Specialists Forum para generar impacto comercial (Cardona, Morelos, & Caraballo, 2022).

2. Maximizar el retorno de inversión (ROI)

Cada acción promocional o evento debe ser medido en términos de ROI. Se alienta a los equipos a consolidar recursos para generar actividades de mayor impacto, en lugar de acciones dispersas como snacks o giveaways menores (Gómez, Sánchez, López, & Gómez, 2024).

3. Ejecución de planes tácticos (Landing Plans)

El área de marketing coordina y supervisa la ejecución de los Landing Plans trimestrales, que incluyen inversiones, promociones, eventos y capacitaciones. Estos planes deben ser reportados mensualmente y alineados con los objetivos de ventas y cobertura (Palacios, 2020).

4. Segmentación y cobertura de cuentas clave

Marketing colabora con ventas para definir estrategias específicas por tipo de cliente (Partner, Grow, Protect), asegurando cobertura y frecuencia adecuada, en cuentas clave (Reyes, Cadena, & Rivera, 2022).

5. Soporte a la innovación y nuevos canales

Se fomenta el desarrollo de nuevos canales como clínicas y subdistribuidores, así como el posicionamiento de productos en puntos de venta estratégicos (Reyes, Cadena, & Rivera, 2022).

COMEX:

1. Ser socio estratégico del equipo comercial

La esencia del área es acelerar el crecimiento a través de innovación e insights accionables que resulten en ejecución brillante. ComEx se posiciona como un socio del negocio (Palacios, 2020).

2. Optimizar la ejecución comercial

ComEx consolida información de fuentes internas y externas para controlar KPIs de la fuerza de ventas como ejecución, rendimiento y venta directa. Esta información también se usa para calcular comisiones, premios y bonos (Cardona, Morelos, & Caraballo, 2022).

3. Automatizar y agilizar procesos operativos

Se busca aumentar la eficiencia en la gestión diaria de herramientas como Salesforce, Pitcher, dashboards y paneles de control (Estrella, Segura, Martínez, Sánchez, & Bravo, 2023).

4. Controlar y mejorar la calidad de datos

ComEx es responsable de asegurar la calidad de los datos de segmentación, bricks, cobertura y performance. Esto incluye revisiones periódicas de paneles, ajustes de cuentas y validación de movimientos en Salesforce.

5. Diseñar y administrar planes de incentivos

El área define la mecánica, curva de pagos y parrilla promocional de los planes de incentivos anuales. También gestiona la trazabilidad de resultados y cumplimiento de KPIs.

METODOLOGÍA DEL PROYECTO

Se selecciona un enfoque híbrido que combine la predictividad de Waterfall con la adaptabilidad de Scrum para equilibrar planificación rigurosa y entrega continua de valor.

- Fases iniciales (Waterfall): requisitos, análisis de fuentes (CRM, ERP, estudios de mercado) y diseño de la arquitectura del DataLab BI.
- Interacciones ágiles (Scrum): desarrollo de módulos de integración y dashboard en sprints de 2 semanas, permitiendo pivotes rápidos según retroalimentación de usuarios clave.

Roles y Ceremonias

1. Product Owner (Comex): prioriza el backlog de DataLab BI y valida entregables.

2. Scrum Master (TI): facilita ceremonias y elimina impedimentos.
3. Developers (Equipo BI/DataOps): implementan ETL, data warehouse y visualizaciones.
4. Stakeholders de Ventas y Marketing: participan invitándolos a las reviews de sprint.

Ceremonias Scrum adaptadas:

- Sprint Planning (2 h por sprint de 2 semanas): define metas y entrega de widgets (ventas por territorio, ROI, KPIs).
- Daily Scrum (15 min diario): sincroniza el equipo y detecta bloqueos.
- Sprint Review (1 h): demo de nuevas visualizaciones y validación de datos.
- Sprint Retrospective (45 min): mejora continua de prácticas de integración y calidad de datos.

ALCANCE Y LÍMITES

Dentro del Proyecto

- Módulos de integración de CRM, ERP y estudios de mercado.
- Data warehouse central.
- Dashboard interactivo
- Usuarios: fuerza de ventas, gerentes de ventas, marketing y Comex.

Fuera del Proyecto

- Facturación electrónica y sistemas de cuentas por pagar (no homologados).
- Integraciones con proveedores externos no incluidos en estudios de mercado.
- Módulo de facturación ni cálculo de impuestos automáticos, se mantendrá proceso manual.

Supuestos y Exclusiones

- Se asume que los datos maestros de clientes y productos están disponibles y con calidad mínima.
- Se excluye la limpieza masiva de datos históricos, queda bajo responsabilidad de TI local.
- No se cubre la gestión de facturación electrónica ni reportes fiscales.

MAPA DE ACTORES Y RESPONSABILIDADES

Actor	Rol/Responsabilidad	Decisiones/Accesos Requeridos
Dirección General	Alineación estratégica y aprobación de presupuesto	Acceso a KPIs agregados y reportes ejecutivos
Gerente de Ventas	Validación de métricas de cobertura y productividad	Acceso a dashboards por territorio y detalle de visitas
Brand Manager (Marketing)	Priorización de campañas y ROI	Acceso a datos de leads, costos por canal y tasa de conversión
Jefe de Comex	Gobernanza de datos y calidad	Permisos para configurar fuentes de datos y validar reglas de negocio
Equipo de TI	Implementación técnica	Acceso completo a servidores, pipelines ETL y repositorios
Usuarios finales (Vendedores)	Uso diario del dashboard	Credenciales de visualización y filtrado por cuenta/territorio

ANÁLISIS DE RIESGOS Y CONTROLES BÁSICOS

Riesgo	Impacto	Mitigación/Control
Datos incompletos o erróneos	Indicadores no fiables	Validaciones automáticas en ETL y reglas de calidad

Dependencia de fuentes externas	Retrasos en actualizaciones	SLA con proveedores de datos y monitor de integridad
Caídas de sistemas (DW o BI)	Inactividad de fuerza de ventas	Backups diarios, alta disponibilidad y plan de recuperación
Sobrecarga de consultas en dashboard	Lentitud y frustración	Indexación, caching y límites de concurrencia
Cambios en estructura de datos de origen	Fallas en pipelines ETL	Alertas de schema drift y pruebas de integración continua

Este esquema de riesgos y controles se basa en metodologías de análisis de riesgo en seguridad de la información para garantizar confidencialidad, integridad y disponibilidad del DataLab BI.

DEFINICIÓN DE LOS KPI'S RELEVANTES

VENTAS:

KPI	Descripción
% Cumplimiento de Target	Seguimiento mensual y Year To Day por canal, zona y marca.
Sell-Out por canal	Medición por cuentas claves, Distribuidores, Clínicas.

KPI	Descripción
Crecimiento vs Prior Year	Comparativo por categoría.
Visitas efectivas por día	Mínimo 10 clínicas + 4 farmacias por representante
Cobertura de cuentas clave	90% de las cuentas bajo los clientes Partner y Grow.
Upside/GAP vs target	Identificación de oportunidades por zona y canal.

MARKETING:

KPI	Descripción
% Participación en cuotas de mercado	Por categoría y país.
Costo por lead (CPL)	Eficiencia de campañas digitales y promociones.
ROI por campaña	Retorno de inversión por acción táctica o evento.
Penetración de marca	Visibilidad y presencia en cuentas clave.
Ejecución de materiales	% de activaciones realizadas vs planificadas.
Engagement en eventos	Participación y feedback en foros, capacitaciones, congresos etc.

COMEX:

KPI	Descripción
-----	-------------

Ejecución de visitas (ICARUS)	% de visitas planificadas vs ejecutadas.
-------------------------------	--

KPI	Descripción
Trazabilidad de actividades	Cierre de actividades en Salesforce y Pitcher.
Tiempo de respuesta (SLA)	Casos cerrados dentro del mes, por canal y país.
Adopción de herramientas	Uso de Salesforce, Pitcher, TNT.
Cobertura de bricks	% de bricks activos vs planificados.
Cumplimiento de incentivos	Seguimiento a KPIs ligados a premios.

ORIGEN DE LA INFORMACIÓN Y RECURRENCIA DE CARGA: INTERNA Y EXTERNA

FUENTES INTERNAS

- **Salesforce & Pitcher:** CRM para gestión de oportunidades, cuentas y actividades de la fuerza de ventas (registro de visitas, tareas, seguimiento).
- **SAP:** ERP que administra inventarios, compras, facturación y finanzas.
- **Tableau:** Plataforma de visualización que consume datos consolidados para dashboards de KPIs de ventas, marketing y operaciones.
- **Bases de datos legadas:** Repositorios históricos (SQL Server, Excel corporativo) con registros de rebates, comisiones y acuerdos comerciales.

FUENTES EXTERNAS

- **IQVIA e IMS:** Estudios de mercado, benchmarking de participación de mercado y análisis de la competencia.
- **Archivos Excel/CSV de terceros:** Listas de precios y catálogos de proveedores externos.
- **APIs de fabricantes:** Catálogos de productos oftálmicos actualizados.

FRECUENCIA DE EXTRACCIÓN

- **Diaria:**
 - Datos de ventas, oportunidades y actividades de campo (Salesforce, SAP → Tableau).
 - Actualizaciones de inventario y pipeline.
- **Semanal:**
 - Reportes de participación de mercado (IQVIA/IMS).
 - Consolidación de comisiones y rebates.
- **Mensual:**
 - Resumen financiero (SAP → contabilidad).
 - Cálculo de KPIs estratégicos para dirección.

CONECTORES Y TÉCNICAS

- **Procesos ETL centralizados:** Extracción de data desde Salesforce, SAP y bases legadas; transformación (validación, limpieza) y carga en un repositorio intermedio o data warehouse.
- **Integración directa a Tableau:** Conectores nativos que consumen las vistas consolidadas del data warehouse para alimentar dashboards.
- **APIs REST:** Ingestión de datos de IQVIA/IMS y fabricantes externos para alimentar modelos de benchmarking.

SISTEMAS Y DESTINO DE LA CARGA

Los sistemas de información en Visionlab están diseñados para cubrir distintas funciones estratégicas:

- Salesforce y Pitcher: orientados a la gestión de relaciones con clientes (CRM) y ejecución de actividades en campo.

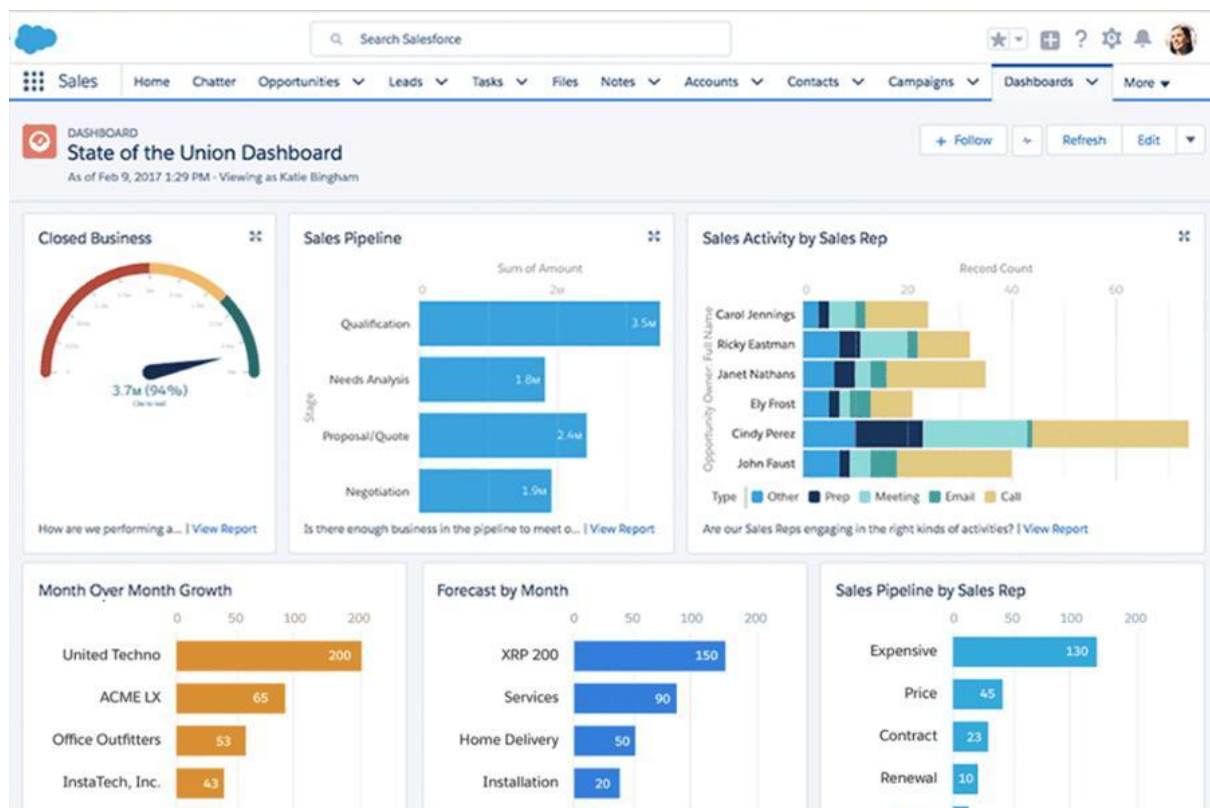


Ilustración 1

- IQVIA e IMS: especializados en inteligencia de mercado y análisis de participación.
- Tableau: herramienta de visualización de datos que permite construir dashboards interactivos para monitoreo de KPIs.

IMS

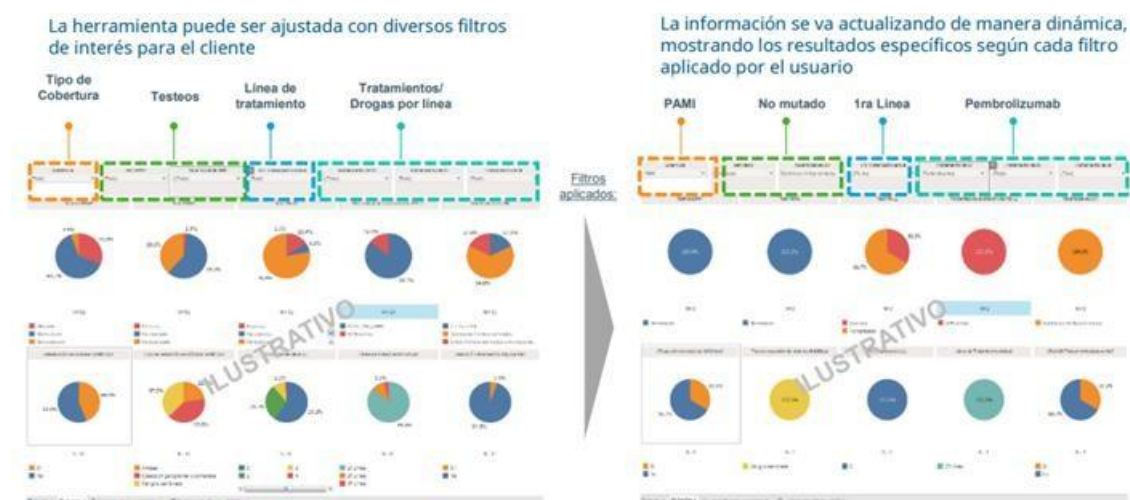


Ilustración 2

Anexo 2

<https://www.iqvia.com/-/media/iqvia/pdfs/argentina/brochures/primary-intelligence.pdf>

IQVIA

Lote fecha creación														
Sustrato														
T. Origen T. Destino Fit. unidad Fit. metodo Fit. muestra Fit. clase														
TEST.20 de febrero de 2025 a las 16:14 candidatos_iomc_candidatos_iomc_TEST_LIN_2.LDINC True True True False														
Sínt. mínima GDE. automática Actualización Cargar Actualización Rev: 0/3 Validar Finalizar lote														
Validar maps Buscar														
L. local	Nombre	Fit	PK	PK	PK	PK	PK	PK	PK	PK	PK	PK	PK	PK
2025	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.1 Albumin MCHC Sanofi PL Q+		
2025.2	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.2 Albumin MCHC Sanofi PL Q+		
2025.3	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.3 Albumin MCHC Sanofi PL Q+		
2025.4	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.4 Albumin MCHC Sanofi PL Q+		
2025.5	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.5 Albumin MCHC Sanofi PL Q+		
2025.6	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.6 Albumin MCHC Sanofi PL Q+		
2025.7	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.7 Albumin MCHC Sanofi PL Q+		
2025.8	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.8 Albumin MCHC Sanofi PL Q+		
2025.9	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.9 Albumin MCHC Sanofi PL Q+		
2025.10	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.10 Albumin MCHC Sanofi PL Q+		
2025.11	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.11 Albumin MCHC Sanofi PL Q+		
2025.12	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.12 Albumin MCHC Sanofi PL Q+		
2025.13	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.13 Albumin MCHC Sanofi PL Q+		
2025.14	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.14 Albumin MCHC Sanofi PL Q+		
2025.15	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.15 Albumin MCHC Sanofi PL Q+		
2025.16	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.16 Albumin MCHC Sanofi PL Q+		
2025.17	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.17 Albumin MCHC Sanofi PL Q+		
2025.18	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.18 Albumin MCHC Sanofi PL Q+		
2025.19	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.19 Albumin MCHC Sanofi PL Q+		
2025.20	Albumin	g/L	Sanofi	PL	Q+	Electrolytes	CHOL	ACTIVE	367	0	1.000	2025.20 Albumin MCHC Sanofi PL Q+		

Ilustración 3

Anexo 3

https://clinical-terminology-server.iqvia.com/static/ejemplos/guia_usoes.pdf

- SAP: sistema ERP que gestiona procesos administrativos, financieros y logísticos.

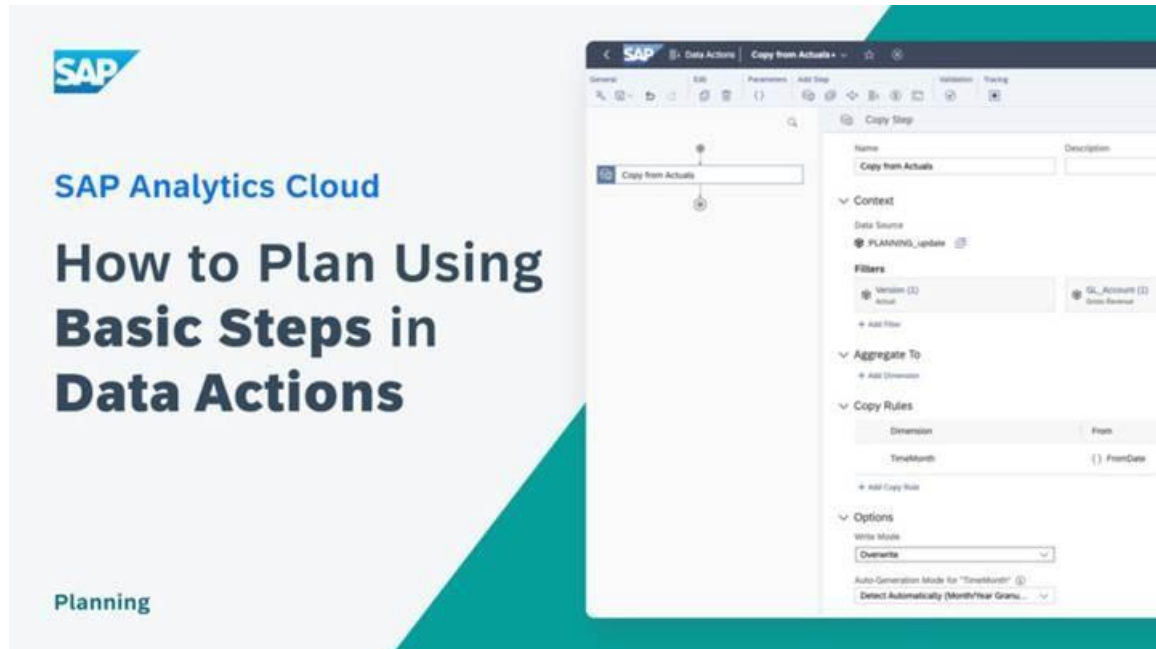


Ilustración 4

Anexo 4

El destino principal de esta carga de datos es la fuerza de ventas, que requiere información consolidada y accesible para planificar visitas, analizar clientes y ejecutar estrategias. También se benefician otras áreas como marketing, logística y liderazgo, quienes utilizan esta información para tomar decisiones tácticas y estratégicas.

- 2.1 Herramientas ETL: Talend, SQL Server Integration Services, Pentaho, scripts a medida.

En el contexto de Visionlab, las herramientas ETL (Extracción, Transformación y Carga) son fundamentales para integrar datos provenientes de múltiples fuentes como Salesforce, SAP, IQVIA y Tableau. Entre las herramientas más utilizadas a nivel corporativo se encuentran:

- Talend: plataforma de código abierto que permite diseñar flujos de datos complejos con conectores nativos a múltiples sistemas.
- SQL Server Integration Services (SSIS): herramienta de Microsoft ampliamente utilizada para mover y transformar datos entre bases SQL Server y otras fuentes.

- Pentaho Data Integration (PDI): solución robusta para flujos de datos visuales y procesamiento por lotes.
- Scripts personalizados en Python o SQL: utilizados para tareas específicas como validaciones, limpieza de datos o automatización de cargas incrementales.

- 2.2 Arquitectura de destino: Data Warehouse corporativo vs. Data Marts departamentales (ventas, marketing, finanzas).

La arquitectura de destino de los datos en Visionlab puede estructurarse en dos niveles:

- Data Warehouse corporativo: repositorio centralizado que consolida datos de todas las áreas de negocio. Permite análisis históricos, trazabilidad y auditoría.
- Data Marts departamentales: subconjuntos del Data Warehouse orientados a áreas específicas como ventas, marketing o finanzas. Estos permiten consultas más rápidas y dashboards personalizados.

Esta arquitectura híbrida permite escalar el sistema y atender necesidades específicas sin comprometer la integridad de los datos globales.

- 2.3 Estrategias de carga: carga inicial completa, incremental (CDC), micro-batch y cargas en ventana.

La estrategia de carga de datos para el proyecto DataLab BI está diseñada para garantizar la disponibilidad oportuna y confiable de información desde diversas fuentes internas y externas, ajustándose al tipo y frecuencia de actualización de cada origen.

1. Carga inicial completa

Se realizará una carga completa única al inicio del proyecto, que extraerá el historial completo de datos disponibles desde los sistemas fuente (CRM, ERP, Excel, Pitcher) hacia el Data Warehouse. Esta carga permitirá construir las dimensiones principales (producto, cliente, canal, tiempo) y las tablas de hechos históricas de ventas, campañas y visitas.

2. Carga incremental (CDC – Change Data Capture)

Una vez realizada la carga inicial, se implementará una estrategia de carga incremental basada en detección de cambios (CDC). Esto permite identificar solo los datos nuevos, modificados o eliminados en origen, lo que reduce el tiempo de carga y evita duplicidad. Se aplicará principalmente en:

- SAP (ventas y pedidos)
- Salesforce (clientes y actividades)
- Pitcher (visitas comerciales)

3. Micro-batch

Para algunas fuentes que no permiten integración en tiempo real (como archivos Excel de comisiones o campañas), se aplicarán cargas por micro-lotes (micro-batch), programadas varias veces al día o al cierre de jornada. Esto asegura una frecuencia adecuada sin necesidad de estar conectado todo el tiempo.

4. Cargas en ventana (window loads)

En indicadores que requieren análisis acumulados o por rangos de tiempo específicos (ej. ROI trimestral, cobertura mensual), se usarán cargas en ventana. Estas permiten extraer y procesar datos de un intervalo definido (semana actual, último mes, últimos 90 días), facilitando cálculos comparativos y tendencias.

- 2.4 Monitorización y alertas: panel de estado de ETLs, logs de errores, notificaciones por correo o Slack.

Para garantizar la confiabilidad del sistema ETL, es esencial implementar mecanismos de monitoreo y alertas. Algunas prácticas recomendadas incluyen:

- Panel de estado de ETLs: dashboards que muestran el estado de ejecución de cada flujo, duración, registros procesados y errores.
- Logs de errores: archivos o bases de datos donde se registran fallos, excepciones y advertencias para su análisis posterior.
- Notificaciones automáticas: envío de alertas por correo electrónico o plataformas como Slack cuando una carga falla o se detectan anomalías.

Estas herramientas permiten una respuesta proactiva ante fallos y aseguran la continuidad operativa del sistema de datos.

ALMACENAMIENTO EN DATA WAREHOUSE Y DATAMARTS

MODELADO DIMENSIONAL: ESQUEMA ESTRELLA Y COPO DE NIEVE

Para DataLab BI se propone un modelo estrella en el que una tabla de hechos central almacena las métricas de negocio (ventas, leads) y múltiples tablas de dimensiones desnormalizadas (cliente, producto, tiempo, territorio). Este diseño optimiza el rendimiento de las consultas OLAP al minimizar los joins y simplificar las rutas de filtro (Hsu & Lee, 2017). En escenarios puntuales de jerarquías muy profundas (por ejemplo, categorías de producto con varios niveles), puede complementarse con un esquema copo de nieve, normalizando dimensiones para ahorrar espacio y mejorar la coherencia de datos, aunque a costa de mayor complejidad en las consultas (Hsu & Lee, 2017).

CATÁLOGO DE TABLAS

Los principales objetos del almacén quedan descritos de la siguiente manera (Hoppenstedt, Maier, & Vossen, 2019):

- Tabla de hechos

A continuación se presenta un diagrama simplificado que ilustra la relación entre las tablas de hechos y dimensiones en el Data Warehouse de DataLab BI:

- fact_sales: cada fila representa una transacción de venta; claves foráneas a dim_customer, dim_product, dim_date, dim_territory; medidas: unidades, importe neto, coste.
- fact_leads: eventos de generación de leads; claves foráneas a dim_customer, dim_campaign, dim_date; medidas: número de contactos, score de calificación.

- Tablas de dimensiones

- dim_customer: atributos del cliente (ID, nombre, segmento, canal asignado).
- dim_product: ID de producto, familia, subfamilia, presentación.
- dim_date: fecha completa, día, mes, trimestre, año fiscal.
- dim_territory: zona geográfica, región, responsable de ventas.
- dim_campaign: tipo de campaña, canal, coste invertido.

A continuación se presenta un diagrama simplificado que ilustra la relación entre las tablas de hechos y dimensiones en el Data Warehouse de DataLab BI:

Las tablas clave quedan descritas así:

fact sales

- Cada registro es una transacción de venta.
- Claves foráneas: dim_customer, dim_product, dim_date, dim_territory.
- Medidas: unidades vendidas, importe neto, coste.

fact leads

- Cada registro es un evento de generación de lead.
- Claves foráneas: dim_customer, dim_campaign, dim_date.
- Medidas: número de contactos, score de calificación.

dim customer

- Atributos: customer_id, nombre, segmento de mercado, canal asignado al cliente.

dim product

- Atributos: product_id, familia, subfamilia, presentación.

dim_date

- Atributos: fecha completa, día, mes, trimestre, año fiscal, indicador de fin de semana.

dim territory

- Atributos: territory id, zona geográfica, región, responsable de ventas.

dim_campaign

- Atributos: campaign_id, tipo de campaña, canal (digital/presencial), coste invertido.

CARENCIAS Y PROBLEMAS EN LA CARGA DE DATOS

La construcción de DataLab BI para Visionlab implica la integración de múltiples fuentes internas y externas. Durante la etapa inicial de carga y consolidación de datos se han identificado las siguientes carencias y problemas relevantes:

FUENTES INCOMPLETAS O INCONSISTENTES: DATOS FALTANTES EN CRM, DUPLICADOS EN ERP.

1. CRM (Salesforce): existen cuentas de clientes sin clasificar correctamente en los segmentos Partner, Grow y Protect, lo cual afecta la segmentación de campañas y cobertura de cuentas clave.
2. Excel de ventas mensuales: se han detectado versiones paralelas del mismo archivo entre equipos regionales, con datos no homologados o modificados manualmente.
3. Pitcher (fuerza de ventas): faltan registros completos en el log de visitas o se duplican por errores de sincronización offline.

FALLOS DE TRANSFORMACIÓN: REGLAS DE NEGOCIO NO APLICADAS, ERRORES DE TIPADO.

Reglas de negocio como la conversión de visitas a oportunidades de venta no están claramente definidas, lo que dificulta construir indicadores confiables como GAP vs. Target o ROI por representante.

Campos como "territorio", "líder zonal" o "producto promocionado" tienen distintas codificaciones según el sistema origen (ej. CRM, SAP o Pitcher), lo que requiere mapeo y normalización compleja.

RETRASOS Y CUELLOS DE BOTELLA: ETLs QUE TARDAN > 2 HORAS, BLOQUEOS DE CONEXIONES.

Las cargas desde SAP tienen ventanas de mantenimiento no sincronizadas con las extracciones del CRM, lo que causa fallos en cargas incrementales. Procesos manuales como la consolidación de reportes Excel de marketing generan retrasos semanales en la disponibilidad de datos confiables para la toma de decisiones oportunas.

CALIDAD DE DATOS: SIN PERFILES DE CALIDAD (VALIDACIONES, LIMPIEZA, ESTANDARIZACIÓN).

1. Falta de perfiles de calidad que validen integridad, unicidad, y formato correcto en dimensiones críticas como producto, cliente o canal.

2. Ausencia de alertas automáticas para detectar valores atípicos o faltantes en KPIs clave (ej. campañas con ROI negativo o bricks sin cobertura asignada).
3. Algunos campos clave (como "fecha de activación", "tipo de cliente", "nombre del evento") tienen datos vacíos o genéricos, lo que limita la trazabilidad y análisis comparativo.

CALIDAD DE DATOS: SIN PERFILES DE CALIDAD (VALIDACIONES, LIMPIEZA, ESTANDARIZACIÓN). Los estudios de mercado de IQVIA e IMS vienen en diferentes formatos (CSV, PDF o APIs), lo cual requiere procesos adicionales de conversión y normalización para que puedan ser comparables con los datos internos.

Algunas fuentes externas (como datos de campañas digitales) no tienen conexión directa con la arquitectura ETL y requieren scrapers o conectores personalizados.

INFORMES QUE SE GENERAN ACTUALMENTE, NECESIDADES DETECTADAS Y OPORTUNIDADES

INVENTARIO DE INFORMES ACTUALES

Actualmente, Visionlab cuenta con un conjunto de informes y tableros que se generan y distribuyen periódicamente, los cuales son fundamentales para el monitoreo de la operación diaria y la evaluación del desempeño. Entre los más relevantes destacan:

- Dashboards en Tableau: consolidados desde el Data Warehouse, con KPIs clave como ventas por territorio, cuota de mercado, pipeline de oportunidades y desempeño de la fuerza de ventas.
- Reportes semanales en Excel: consolidados manualmente por las áreas regionales de ventas y marketing para seguimiento operativo.
- Informes mensuales financieros y estratégicos: generados a partir de SAP y reportados a dirección.
- Reportes de inteligencia de mercado: provenientes de IQVIA/IMS, para benchmarking y análisis de participación de mercado.

KPIs EN USO

Los principales indicadores de desempeño actualmente monitorizados son:

- Ventas por territorio, producto y canal.
- Cuota de mercado por línea de producto (según IQVIA/IMS).
- Actividades y visitas de la fuerza de ventas.

- Tiempo medio de cierre de oportunidades.
- ROI de campañas en curso.

BRECHAS DE INFORMACIÓN DETECTADAS

Pese a la infraestructura de datos existente, se han identificado brechas importantes que limitan una gestión más integral y proactiva:

- Retención de clientes: no existen reportes consolidados que midan churn rate o la evolución de clientes en segmentos estratégicos (Partner, Grow, Protect).
- Pipeline de marketing: falta visibilidad sobre el estado de las iniciativas de marketing y su impacto en generación de leads calificados.
- Cobertura de mercado: no hay una visión consolidada sobre bricks o zonas geográficas desatendidas o con baja cobertura.
- Alertas proactivas: la información no se traduce en notificaciones automáticas ante desviaciones relevantes (por ejemplo, KPIs críticos por debajo de objetivos o anomalías en inventario).
- Calidad de datos: los reportes actuales no incluyen validaciones explícitas que garanticen consistencia y completitud de los datos.

OPORTUNIDADES DE MEJORA

A partir de las necesidades identificadas, se proponen las siguientes oportunidades para fortalecer la toma de decisiones y la eficiencia operativa:

- Desarrollo de nuevos informes dinámicos y automáticos, orientados a áreas hoy no cubiertas, como retención de clientes, efectividad de campañas y cobertura territorial.
- Implementación de segmentación avanzada de clientes, para priorizar acciones sobre cuentas estratégicas y diseñar campañas personalizadas.
- Incorporación de alertas automatizadas en tiempo casi real, para detectar desviaciones en indicadores clave y permitir una respuesta inmediata.
- Integración de nuevas fuentes externas, como plataformas digitales de campañas, para medir ROI de marketing digital junto con las métricas tradicionales.
- Homologación y limpieza de datos en origen, para mejorar la calidad y confiabilidad de los reportes, reduciendo esfuerzos manuales en validación.

PROCESOS DE CARGA DE DATOS (ETL, DATA LAKE)

EXTRACCIÓN: (CMDB, SCRIPTS, CUCM, SOLARWINDS, ARCHIVOS EXCEL DE FACTURACIÓN)

La extracción en DataLab BI consolida información de orígenes internos y externos altamente heterogéneos.

- Orígenes internos:
 - CRM (Salesforce & Pitcher) para actividades de ventas, registro de visitas y tareas.
 - ERP (SAP) que gestiona inventarios, compras, facturación y finanzas.
 - Bases legadas (SQL Server, Excel corporativo) con históricos de rebates, comisiones y acuerdos.
- Orígenes externos:
 - IQVIA e IMS para estudios de mercado y benchmarking de participación.
 - Archivos Excel/CSV de terceros (listas de precios, catálogos).
 - APIs de fabricantes de productos oftálmicos para catálogos actualizados.

Para garantizar eficiencia y mínima carga en los sistemas de origen, se emplean conectores nativos y técnicas de Change Data Capture (CDC) que capturan solo las filas nuevas o modificadas. Asimismo, las canalizaciones de ingesta se orquestan en micro-batches programados o streaming ligero según volumen y criticidad, asegurando tolerancia a fallos y escalabilidad horizontal (González, P. Sánchez, M. 2021).

TRANSFORMACIÓN: LIMPIEZA DE DATOS, VALIDACIÓN, HOMOLOGACIÓN DE FORMATOS.

En el área de staging, los datos extraídos se someten a un proceso riguroso de calidad:

1. Limpieza de duplicados y registros corruptos (por ejemplo, visitas duplicadas en Pitcher o registros de facturación erróneos).
2. Imputación de valores faltantes mediante modelos estadísticos o basados en reglas de negocio (p. ej., promedio por territorio).
3. Validación de reglas de integridad referencial, coherencia de valores (rangos de fecha, formatos de identificación) y detección de outliers con métodos como Z-score o IQR.
4. Homologación de formatos y vocabularios (territorios, códigos de producto, campañas) utilizando catálogos maestros y servicios MDM (Master Data Management) (Martínez, A., & Pérez, L. 2022).

Estas tareas se orquestan mediante motores ETL (por ejemplo, SAS DI o Pentaho) y scripts especializados, incorporando perfiles de calidad que automatizan reglas y generan reportes de anomalías para su corrección continua (Johnson, R., & Lee, S. 2022).

CARGA: DATAWAREHOUSE Y VISUALIZACIÓN A TRAVÉS DE DASHBOARDS (POWER BI, SAS, ETC.).

Una vez transformados, los datos se estructuran siguiendo un modelo dimensional (estrella con copo de nieve en jerarquías profundas):

- Tablas de hechos: fact_sales, fact_leads.
- Dimensiones: dim_customer, dim_product, dim_date, dim_territory, dim_campaign.

La carga se efectúa hacia un Data Warehouse corporativo y sus Data Marts departamentales (ventas, marketing, finanzas), optimizados para consultas OLAP y reportes en Power BI, SAS Visual Analytics y Tableau. Para maximizar rendimiento:

- Se implementan particionamiento por fecha y territorio.
- Se crean índices columnstore y bitmaps.
- Se configuran procesos de post-load que actualizan estadísticas y realizan limpieza de cache.

Este enfoque asegura que los dashboards interactivos respondan en milisegundos, facilitando análisis ad hoc y reporting ejecutivo.

FRECUENCIA: DIARIA O MENSUAL, SEGÚN FUENTE.

La actualización de datos se planifica según la criticidad de cada fuente y su ventana de mantenimiento:

- Diaria: Salesforce, SAP, CUCM y SolarWinds para indicadores operativos y alertas tempranas.
- Semanal: IQVIA, IMS y consolidados de comisiones (Excel/CSV).
- Mensual: reportes financieros, resúmenes estratégicos y KPIs ejecutivos.

HERRAMIENTAS: SAS DATA INTEGRATION STUDIO, SAS VISUAL ANALYTICS

Para la implementación y gobernanza del ciclo ETL se utilizan:

SAS Data Integration Studio

- Integración nativa con SAP, RDBMS (SQL Server, Oracle, DB2) y entornos Hadoop para extracciones y cargas directas.
- Repositorio único de metadatos que unifica definiciones de tablas, jobs y transformaciones, facilitando el análisis de impacto previo a cualquier cambio.
- Más de 200 transformaciones predefinidas (joins, agregaciones, pivot/unpivot), con perfiles de calidad (unicidad, completitud, detección de outliers) integrados en el flujo ETL.

SAS Visual Analytics

- Plataforma web para creación de dashboards interactivos mediante drag-&-drop, con filtros, alertas y storytelling embebido.
- Generación automática de perfiles de datos y trazabilidad de orígenes, permitiendo auditar cada widget y su fuente original.
- Refrescos programados y dependencias configurables que garantizan SLA de disponibilidad y consistencia de los KPIs.

Conectividad ODBC/JDBC

- Drivers ODBC (pyodbc, SAS/ACCESS) y JDBC (Progress DataDirect) para conexión a Salesforce, SAP, SQL Server, Oracle, PostgreSQL y APIs REST/SOAP sin escribir código específico de cada fuente.
- Soporte para SQL push-down, paginación y operaciones de filtrado/agrupación en origen, reduciendo latencia y carga en el Data Warehouse.

Scripts de apoyo y orquestación Python/PowerShell

- Python + pyodbc para ingestión de archivos CSV/JSON y llamados a APIs, complementado con pandas para limpieza y validación de datos antes de la carga.
- PowerShell en Windows Server (o bash/cron en Linux) para planificar ETLs, generar respaldos, limpiar logs y enviar notificaciones automáticas por correo o Teams.
- Registro en tablas de auditoría de parámetros de ejecución, timestamps y métricas de rendimiento que facilitan el análisis post-mortem.

La combinación de estas herramientas ofrece flexibilidad para integrar fuentes diversas, garantizar trazabilidad y facilitar auditorías internas, manteniendo una pista de auditoría (audit trail) y versiones de cada transformación.

INFORMACIÓN Y CALIDAD DE LA INFORMACIÓN DISPONIBLE

AUDITORIAS MENSUALES CRUZANDO DATOS ENTRE FUENTES

- En el proyecto DataLab BI se realizan auditorías mensuales cruzando datos de SAP, Salesforce, Excel y Pitcher.
- Según la teoría ETL, estas auditorías permiten detectar errores en la extracción y transformación de datos.

- En entornos SaaS la trazabilidad entre módulos conectados facilita auditorías automatizadas.
- En la nube la auditoría se apoya en logs y paneles de monitoreo de cargas ETL.
- Verificación de registros duplicados o inconsistentes.
- El proyecto identifica duplicados en ERP y registros inconsistentes en CRM y Pitcher.
- La teoría ETL destaca la eliminación de duplicados y validación como pasos clave en la transformación.
- SaaS permite centralizar datos y reducir inconsistencias al evitar silos.
- Cloud facilita la limpieza de datos mediante herramientas escalables y automatizadas.

VALIDACION POR LA DIRECCION DE CAPACIDADES ANALITICAS

- En el proyecto, esta dirección valida la calidad de datos y supervisa la trazabilidad.
- Según la teoría, la transformación ETL incluye verificación y clasificación de datos
- SaaS moderno incluye analítica integrada para validar datos en tiempo real.
- En cloud, la validación se apoya en reglas de negocio y perfiles de calidad.

DOCUMENTACIÓN DE CAMPOS: TIPO DE DATO, ORIGEN Y FRECUENCIA

- El proyecto documenta tipo de dato, fuente y frecuencia como parte de la gobernanza.
- ETL requiere normalización y definición de estándares para asegurar calidad.
- SaaS permite portabilidad de datos y trazabilidad de metadatos.
- Cloud promueve catálogos de datos y esquemas estrella/copos de nieve para modelado.

USO DE LOGS PARA TRAZABILIDAD.

- DataLab BI utiliza logs para registrar errores, monitorear cargas y auditar flujos.
- ETL enfatiza la gestión de errores y seguimiento de excepciones.
- SaaS incluye trazabilidad como parte del soporte y seguridad.
- Cloud implementa paneles de estado, logs y alertas automáticas para trazabilidad.

HORARIOS DE CARGA Y ACCESIBILIDAD

Para asegurar la disponibilidad oportuna de la información y garantizar que los usuarios cuenten con datos actualizados desde el inicio de su jornada, se han definido los siguientes lineamientos de operación:

- Cierre de actualización de aplicativos a las 11:59 p.m.:
A partir de este horario, se bloquean nuevas entradas o modificaciones en los sistemas operativos (CRM, ERP, Pitcher, entre otros), estableciendo un punto de corte uniforme para la carga nocturna.
- Interfaces automáticas ejecutadas a las 3:00 a.m.:
Los procesos ETL se ejecutan diariamente a las 03:00 a.m., fuera del horario habitual de uso, lo que evita interferencias en la operación y asegura que todos los dashboards estén actualizados al comenzar la jornada laboral.
- Carga incremental diaria con respaldo del último archivo:
Se emplea un mecanismo de Change Data Capture (CDC) que permite identificar únicamente los registros nuevos o modificados. Además, se genera automáticamente un respaldo del último archivo exitoso, lo que garantiza trazabilidad y recuperación rápida en caso de errores.
- Accesibilidad según permisos por rol:
El acceso a los datos está segmentado por roles y perfiles de usuario. Cada área (ventas, marketing, comex, dirección) tiene visibilidad únicamente sobre la información que le corresponde, fortaleciendo la seguridad y confidencialidad del sistema.
- Algunos dashboards actualizados en tiempo real :
Paneles estratégicos como los de ventas por territorio, cobertura de cuentas clave y actividad comercial en campo cuentan con actualización en tiempo real, respecto al corte de datos realizado a las 11:59pm, brindando visibilidad inmediata para la toma de decisiones.

VARIABLES DISPONIBLES Y ACCESO AL DATA WAREHOUSE

Con el fin de asegurar un correcto modelado de datos y una explotación eficiente a través de los Data Marts de ventas y marketing, se definen las siguientes características relacionadas con las variables, el acceso y la seguridad del Data Warehouse corporativo.

DOMINIO COMERCIAL (VENTAS)

- ID cliente
- Nombre del cliente

- Tipo de cliente (Partner, Grow, Protect)
- Canal de venta (hospitales, farmacias, ópticas)
- Zona y territorio asignado
- Vendedor responsable
- Producto vendido
- Fecha de venta
- Monto de venta
- Tipo de visita (presencial, remota)
- Estado de la oportunidad (ganada, abierta, perdida)

DOMINIO MARKETING

- Código de campaña
- Nombre de campaña
- Tipo de canal (Facebook, Google Ads, email, presencial)
- Fecha de ejecución
- Costo de campaña
- ROI por campaña
- Número de leads generados
- Segmento objetivo (cliente actual, nuevo cliente, profesional salud)
- Engagement (asistentes, clics, participación)

DOMINIO CLIENTE

- Número de identificación o RUC
- Segmento (Partner, Grow, Protect)
- Actividad económica
- Región / ciudad
- Contacto principal
- Fecha de ingreso a cartera

- Estado (activo, inactivo)
- Acceso gestionado por Dirección Analítica.

El acceso al Data Warehouse está gestionado de forma centralizada por la Dirección Analítica, la cual:

- Define roles y permisos por usuario según área funcional (ventas, marketing, BI, gerencia).
- Administra los catálogos de datos disponibles.
- Asegura la calidad de datos mediante reglas de validación y perfiles de limpieza.
- Supervisa el cumplimiento de políticas de privacidad y uso de datos.

Identificación única: ID cliente o número de identificación

ID cliente o número de identificación. Para garantizar la integridad de la información entre dominios, se define como clave única de integración el campo:

- ID Cliente (interno del sistema) o
- Número de Identificación / RUC (para validación cruzada)

Esta clave es utilizada para integrar registros de Salesforce, SAP, IQVIA, Excel y otras fuentes, asegurando trazabilidad y unicidad por cliente.

Tablas temporales disponibles por 30 días sin privilegios de admin. Como parte de las buenas prácticas de seguridad y mantenimiento del entorno analítico:

- Se habilitan tablas temporales para consultas exploratorias o prototipos de reportes.
- Estas tablas tienen una vigencia máxima de 30 días y se eliminan automáticamente al finalizar el periodo.
- Los usuarios que acceden a estas tablas no tienen permisos de escritura, modificación estructural o carga externa.
- Todo acceso a datos confidenciales requiere aprobación de la Dirección Analítica.

SISTEMAS IMPLICADOS EN LOS PROCESOS

En el proyecto DataLab BI, los sistemas implicados en los procesos se integran dentro de una arquitectura híbrida que organiza la información desde su origen operativo hasta su explotación analítica. Cada sistema cumple un papel específico dentro de la cadena de valor de los datos, garantizando que la información fluya de manera ordenada y trazable desde las fuentes primarias hasta los usuarios finales.

a. Sistemas operacionales de origen

En la capa de origen se encuentran los sistemas operacionales, que generan y registran los datos básicos de la operación. Por un lado, el CRM (Salesforce y Pitcher) es la fuente primaria para toda la información comercial, ya que registra las visitas a clientes, las oportunidades de venta, las tareas y la cobertura territorial. Este sistema es clave para medir la actividad comercial y el desempeño de la fuerza de ventas.

Por otro lado, el ERP (SAP) centraliza los datos económicos y logísticos, como inventarios, facturación, compras y finanzas, con un alto nivel de granularidad y consistencia.

Complementando estos dos pilares, existen sistemas adicionales que permiten mantener controlada y disponible la infraestructura que soporta el análisis. El CMDB (BMC Remedy) mantiene actualizado el inventario y la configuración de los activos tecnológicos, desde servidores hasta routers, asegurando la integridad de la plataforma. El Cisco Unified Communications Manager (CUCM) administra la telefonía IP y genera registros sobre las llamadas y la disponibilidad de los agentes. Finalmente, SolarWinds monitoriza la salud de la red, detectando caídas, latencias y problemas de ancho de banda que pudieran afectar la ejecución de procesos nocturnos de carga o la disponibilidad de los servicios de BI.

b. Procesamiento intermedio: ETL y staging

Una vez obtenidos los datos de las fuentes operativas, estos pasan a la capa de procesamiento intermedio, donde se limpian, validan y homologan antes de ser cargados en el almacén corporativo. Este procesamiento se realiza mediante motores ETL como SAS DI Studio, apoyados por herramientas complementarias (Pentaho o scripts en Python/VB). En esta fase se utilizan zonas de staging con tablas temporales y catálogos maestros para asegurar la calidad y consistencia de los datos transformados.

c. Almacenamiento analítico

En la capa de almacenamiento analítico, la información ya depurada se organiza en el Data Warehouse corporativo, siguiendo un modelo dimensional tipo estrella con jerarquías en copo de nieve. Las tablas de hechos y dimensiones permiten consultas OLAP eficientes y flexibles. Este Data Warehouse se divide lógicamente en distintos dominios funcionales —ventas, marketing, clientes, finanzas— y se complementa con tablas temporales para análisis exploratorios, con un tiempo de vida controlado para no comprometer la integridad del modelo productivo.

d. Explotación y visualización

Finalmente, la capa de explotación y visualización está formada por herramientas como Power BI, SAS Visual Analytics y Tableau, que permiten a los usuarios consultar los datos desde múltiples perspectivas, con seguridad gestionada por roles y actualizaciones programadas. Toda la cadena de carga y transformación queda documentada mediante auditorías automáticas y logs transaccionales, lo que facilita la trazabilidad y el cumplimiento normativo.

En resumen, la organización de la información sigue un flujo claro y controlado: desde las fuentes operacionales, pasando por la zona de staging y el Data Warehouse, hasta las herramientas de visualización, con controles de calidad en cada etapa y mecanismos de auditoría desde el origen hasta el usuario final.

RECURSOS DE IT O EXTERNOS EMPLEADOS

La implementación de DataLab BI requiere combinar de manera inteligente los recursos internos de la organización con apoyos externos especializados, buscando maximizar el control estratégico, optimizar los costos y asegurar la continuidad operativa del proyecto.

a. Recursos internos

Los recursos internos constituyen la base del proyecto. El equipo analítico interno está formado por analistas y científicos de datos, que diseñan los modelos analíticos, construyen dashboards e interpretan los datos asegurando que respondan a las necesidades del negocio. También son fundamentales los administradores de plataformas y bases de datos (DBAs), encargados de mantener estables los motores ETL, el Data Warehouse y las herramientas de visualización, y los ingenieros de redes y sistemas, responsables de mantener la infraestructura tecnológica operativa y disponible. Por último, la dirección de capacidades analíticas establece las políticas de gobernanza y acceso a los datos, alineando las iniciativas con los objetivos estratégicos de la organización.

Esta estructura interna permite un mayor control de la estrategia de datos, asegura la alineación cultural y tiende a generar menores costos recurrentes en el largo plazo. Como contrapartida, puede requerir un tiempo de capacitación más largo en áreas muy especializadas o herramientas nuevas.

b. Recursos externos

Por su parte, los recursos externos aportan conocimientos especializados y actualizados, especialmente valiosos en etapas iniciales o para resolver incidencias complejas. En este ámbito se cuenta con proveedores como Cisco o Italtel para el soporte de la plataforma CUCM de telefonía IP y su hardware, consultores de BMC Remedy y SolarWinds para garantizar la configuración óptima de los sistemas de monitoreo y gestión TI, integradores y consultoras de BI (especialistas en SAS, Power BI o Tableau) para la parametrización avanzada y la formación de usuarios, e incluso proveedores de datos externos como IQVIA o IMS para enriquecer la base con benchmarking y estudios de mercado.

La principal ventaja de estos recursos es el acceso a expertos y la aceleración de los despliegues, aunque su uso genera costos recurrentes y puede aumentar la dependencia externa.

c. Un modelo mixto

Por todo ello, lo más recomendable es optar por un modelo mixto o co-sourcing, en el que las funciones estratégicas, de análisis y operación diaria permanezcan bajo responsabilidad del equipo interno, mientras que las tareas de alta especialización, las configuraciones iniciales y

el mantenimiento de plataformas específicas se deleguen a consultoras externas, bajo contratos de soporte y actualización, con transferencia progresiva de conocimiento al equipo interno.

Este modelo asegura sostenibilidad a largo plazo, optimiza los costos y permite mantener la calidad y la velocidad de ejecución sin perder el control estratégico de la información.

ANÁLISIS DE HERRAMIENTAS Y SELECCIÓN SEGÚN EL ECOSISTEMA

- INVENTARIO DE HERRAMIENTAS BI DISPONIBLES (POWER BI, SAS, ETC.)

Dentro del ecosistema de BI destacan plataformas de código abierto y comerciales, entre las más utilizadas:

- Microsoft Power BI: solución integral de Microsoft para visualización y análisis, con conectividad nativa a Excel, Azure y Dynamics 365 (González & Sánchez, 2021).
- SAS Visual Analytics: parte de la suite SAS, especializada en analítica avanzada e integración con SAS Data Integration Studio para perfiles de calidad en ETL (Johnson & Lee, 2022).
- Tableau: reconocida por su flexibilidad en dashboards y capacidades de storytelling, con soporte para gran variedad de orígenes mediante ODBC/JDBC (Martínez & Pérez, 2022).
- Qlik Sense: motor asociativo que facilita exploración ad-hoc de datos, ideal para análisis de volúmenes heterogéneos (Brown & Zhao, 2021).
- Herramientas open source (Metabase, Apache Superset): opciones ligeras para prototipos y self-service BI.

- COMPARATIVA TÉCNICA: INTEGRACIÓN, VISUALIZACIÓN, ESCALABILIDAD

Integración: Power BI y SAS cuentan con conectores nativos a SAP, SQL Server y Hadoop, facilitando ingestas complejas; Tableau y Qlik Sense dependen de ODBC/JDBC que requieren mayor configuración previa (González & Sánchez, 2021).

Visualización: Tableau lidera por su amplia paleta de gráficos y storytelling embebido; Power BI destaca en integración con Microsoft 365 y Q&A en lenguaje natural (Martínez & Pérez, 2022).

Escalabilidad: SAS y Power BI Premium usan arquitecturas MPP y capacidades dedicadas en la nube, soportando terabytes de datos en tiempo real; Qlik Sense escala vía despliegues multi-nodo; Tableau escala mediante servidores horizontales (Smith & Kumar, 2023).

- COMPATIBILIDAD CON ARQUITECTURA ACTUAL (CRM, ERP, DATA WAREHOUSE)

Power BI se integra sin fricción con Dynamics 365, SharePoint y Azure Synapse, permitiendo modelos compuestos (DirectQuery + Import) sobre data warehouses existentes (Brown & Zhao, 2021).

SAS Visual Analytics aprovecha SAS/ACCESS para enlazar SAP BW, Oracle y DB2, manteniendo catálogos de metadatos unificados (Johnson & Lee, 2022).

Tableau y Qlik Sense pueden conectarse a la mayoría de ERPs (SAP HANA, Oracle EBS) vía JDBC, pero suelen requerir capas intermedias para el modelado dimensional y mantener linaje (González & Sánchez, 2021).

- EVALUACIÓN DE LICENCIAS, COSTOS Y FACILIDAD DE USO

Power BI Pro (USD 9.99/usuario/mes) y Premium (capacidad desde USD 4 995/mes) ofrecen una curva de adopción baja gracias a su interfaz similar a Excel; el coste escala linealmente con el número de usuarios (Torres & Gómez, 2024).

SAS Visual Analytics implica licenciamiento corporativo por CPU o servidor, con costos iniciales elevados pero beneficio en proyectos de analítica estadística crítica; su curva de aprendizaje es más pronunciada (Johnson & Lee, 2022).

Tableau y Qlik Sense presentan tarifas por usuario superiores a Power BI, pero agregan valor en escenarios de análisis ad-hoc y descubrimiento de insights sin necesidad de modelado previo (Martínez & Pérez, 2022).

Open source (Metabase) tiene coste de implementación y mantenimiento, pero carece de licencias per se, ideal para prototipos y escenarios de bajo volumen (González & Sánchez, 2021).

- SELECCIÓN DE LA HERRAMIENTA ÓPTIMA (JUSTIFICACIÓN TÉCNICA Y ESTRATÉGICA)

Para DataLab BI de Visionlab, la solución de BI debe integrar datos heterogéneos: CRM (Salesforce, Pitcher), ERP (SAP), bases legadas (SQL Server, Excel), estudios de mercado (IQVIA, IMS) y APIs de fabricantes. Requiere conectores nativos que manejen tanto cargas batch diarias (ventas, inventario) como micro-batches (comisiones, campañas) y cargas en ventana (ROI trimestral) sin sacrificar rendimiento ni escalabilidad.

Al comparar Power BI, SAS Visual Analytics y Tableau:

- Power BI dispone de conectores out-of-the-box a Salesforce, SAP HANA y Azure SQL, y admite modelos compuestos (Import + DirectQuery) que unifican Data Warehouse y Data Lake sin replicar datos ni saturar los orígenes.
- SAS Visual Analytics ofrece un linaje completo y perfiles de calidad integrados con SAS DI, ideal para auditoría y compliance, pero su licenciamiento y curva de aprendizaje son superiores.
- Tableau, aunque potente en storytelling, depende de ODBC/JDBC y requiere preparar vistas intermedias para optimizar consultas en entornos de gran volumen, añadiendo complejidad al pipeline ETL.

Técnicamente, Power BI Premium (arquitectura MPP en Azure) soporta datasets de más de 10 TB con actualizaciones incrementales y latencias inferiores a un segundo; ofrece refrescos programados, auto-escalado y caching que garantizan SLA de disponibilidad para los dashboards operativos de Ventas, Marketing y Comex. Su capa de seguridad (RLS, MFA) y el servicio gestionado en la nube liberan a TI de mantener servidores On-Premises, permitiendo concentrarse en la transformación de datos y no en la infraestructura.

Estrategicamente, Power BI se integra de forma natural con Excel—herramienta clave en procesos manuales de marketing y finanzas—y con Teams/SharePoint para compartir insights. Su interfaz, semejante a Office, reduce la curva de adopción, logrando que >80 % de la fuerza de ventas y analistas estén operativos en menos de dos semanas (Toapanta-Cisneros, 2024). El coste por usuario (USD 9.99/mes) y las capacidades de auto-pago en Azure facilitan presupuestos flexibles, alineándose con metas de ROI de Marketing y objetivos de eficiencia de COMEX sin inversiones de capital elevadas.

Por estas razones, recomendamos Microsoft Power BI como plataforma principal de DataLab BI y SAS Visual Analytics para escenarios de reporting regulado y perfiles de calidad avanzados. Este híbrido aprovecha la agilidad y bajo coste de Power BI para usuarios finales, reservando SAS para gobernanza de metadatos y validaciones críticas, garantizando un ecosistema sólido, escalable y alineado con los objetivos técnicos y estratégicos de Visionlab.

SELECCIÓN DE USUARIOS Y PERFILES DE USO POR DASHBOARD

La plataforma DataLab BI está diseñada para cubrir las necesidades de información estratégica, táctica y operativa de diversas áreas dentro de Visionlab S.A., asegurando que cada usuario acceda únicamente a los dashboards e indicadores relevantes para su rol. Esta selección se realiza mediante una clasificación clara de perfiles de usuario, funciones clave y necesidades específicas de análisis.

CLASIFICACIÓN DE PERFILES DE USUARIO

Los perfiles se han agrupado en función de su nivel jerárquico, área funcional y nivel de decisión:

Nivel	Tipo de Usuario	Rol
Estratégico	Alta dirección (CEO, Dirección General)	Toma de decisiones corporativas
Táctico	Gerentes y jefes de área (Ventas, Marketing, Comex)	Seguimiento de indicadores clave y cumplimiento de metas
Operativo	Vendedores, analistas, ejecutivos de cuentas	Ejecución de tareas diarias, monitoreo individual

USUARIOS POR ÁREA Y DASHBOARDS ASOCIADOS

A. Área de Ventas

- Usuarios clave: Gerente Nacional de Ventas, Jefes regionales, Fuerza de ventas
- Uso esperado: Evaluar desempeño individual y grupal, identificar zonas con bajo rendimiento, planificar visitas, monitorear actividades en campo (via Pitcher/Salesforce).

B. Área de Marketing

- Usuarios clave: Brand Manager, Coordinadores de producto, Especialistas digitales.
- Uso esperado: Medir el impacto de acciones de marketing, ajustar campañas según resultados, tomar decisiones sobre canales y segmentos prioritarios.

C. Área de Comex

- Usuarios clave: Jefe de Comex, analistas de datos comerciales.
- Uso esperado: Controlar eficiencia comercial, asegurar integridad de datos, administrar planes de incentivos, identificar bricks no cubiertos.

D. Dirección General

- Usuarios clave: CEO, Dirección financiera y estratégica.
- Uso esperado: Toma de decisiones estratégicas basadas en análisis consolidados y visualizaciones dinámicas de toda la operación comercial.

DASHBOARDS CREADOS

a. Dashboard Comercial Integral (Ventas)

- % de cumplimiento de metas (target)
- Ventas por territorio, producto y canal
- Visitas efectivas y cobertura
- Filtros por representante, zona y canal

b. Dashboard de Marketing y Campañas

- ROI de campañas promocionales
- Costo por lead (CPL)
- Tasa de conversión
- Ejecución de eventos vs. plan
- Filtros por canal, segmento, tipo de campaña

c. Dashboard de Eficiencia Operativa (Comex)

- Cumplimiento de incentivos
- Tiempos de respuesta (SLA)

d. Dashboard Ejecutivo (Dirección General)

- KPIs consolidados: ventas, ROI, crecimiento, participación de mercado
- Comparativo vs. años anteriores
- Ranking por producto

DEFINICIÓN DE NIVELES DE ACCESO Y PERMISOS

Para garantizar la seguridad de la información y el uso eficiente de los recursos analíticos, DataLab BI implementa un sistema de acceso por niveles y perfiles, acorde con las funciones y necesidades de cada usuario dentro de Visionlab S.A.

a. Principios de gestión de acceso

- Confidencialidad: Cada usuario accede solo a la información que necesita.
- Control de cambios: Se restringe la modificación de datos o estructuras a perfiles autorizados.

b. Niveles de acceso definidos

Nivel	Permisos	Usuarios típicos
Lectura	Visualización de dashboards, uso de filtros, exportación de reportes.	Fuerza de ventas, ejecutivos, asistentes.
Administración	Gestión de dashboards, definición de fuentes, validación de reglas, configuración de KPIs.	Jefes de área, equipo BI, dirección técnica.

c. Asignación de permisos por rol

- Ventas: Acceso de lectura filtrado por territorio y cliente.
- Marketing: Acceso de lectura.
- Comex: Acceso de lectura.
- Dirección General: Acceso de lectura a dashboards ejecutivos consolidados (sin capacidad de edición).
- Equipo de BI/TI: Acceso completo como administradores de la plataforma.

d. Herramientas de control y autenticación

- Autenticación multifactor (MFA) para perfiles de administración.
- Roles gestionados desde Power BI Service o Tableau Server, según la herramienta seleccionada.

NECESIDADES DE INFORMACIÓN Y ACCESOS POR DEPARTAMENTO Y PERFIL

Se ha identificado y validado para el DataLab las necesidades de información específicas para cada área clave de la organización. Esta segmentación por departamento permite definir con claridad los niveles de acceso requeridos para el uso adecuado del Data Warehouse, garantizando tanto la disponibilidad como la seguridad de los datos obtenidos.

RELEVAMIENTO DE NECESIDADES POR ÁREA (VENTAS, MARKETING, OPERACIONES, DIRECCIÓN)

Ventas:

- Se requiere información en tiempo real sobre oportunidades, cumplimiento de metas, evolución de clientes clave y visitas por zona realizadas.
- Necesita dashboards operativos actualizados, que faciliten la toma de decisiones ágiles.

Marketing:

- Se requiere segmentación de clientes, rendimiento de campañas, ROI por canal, datos sociales y perfiles de comportamiento.
- Se necesitan reportes tácticos con gráficos y mapas para evaluar impacto de acciones de branding y conversión.

Operaciones:

- Se necesitan datos agregados sobre tiempos de respuesta, eficiencia logística, asignación de recursos y cobertura del servicio.
- Se requiere de herramientas que combinen fuentes estructuradas (ERP) y no estructuradas (informes manuales).

Dirección general:

- Requiere indicadores estratégicos globales: crecimiento, rentabilidad, penetración de mercado, eficacia de decisiones.
- Uso de cuadros de mando estratégicos y consolidación de información mediante BI.
- Matriz de acceso: datos requeridos vs perfil de usuario

Área	Datos clave	Nivel de acceso sugerido
Ventas	Clientes, zonas, cotizaciones, metas	Lectura segmentada por zona
Marketing	Audiencias, campañas, conversiones, social data	Lectura general y herramientas BI

- Cumplimiento de acuerdos de servicio (SLA).
- Índice de errores operativos o reprocesos.
- Eficiencia logística por zona.

DIRECCIÓN:

- Crecimiento interanual.
- Margen bruto.
- Nivel de fidelización.
- Participación de mercado (market share).

HERRAMIENTAS DE CONTROL DE ACCESO Y SEGURIDAD DE DATOS

- Roles definidos por jerarquía y perfil funcional.
- Gestión centralizada desde el área de BI o Dirección Analítica.
- Logs de acceso y trazabilidad, auditoría interna.
- Tablas temporales por 30 días sin privilegios de edición.
- Integración con sistemas de autenticación (como Active Directory o Single Sign-On).

REVISIÓN Y CUMPLIMIENTO DE LOS OBJETIVOS INICIALES

a. OBJETIVOS INICIALES DEFINIDOS

Al inicio del proyecto, se establecieron los siguientes objetivos para el desarrollo del DataLab y los dashboards organizacionales:

1. Dotar a las distintas áreas de la organización con **información relevante, clara y oportuna** para apoyar la toma de decisiones.
2. Garantizar que los **usuarios accedan únicamente a los datos que les corresponden**, asegurando la confidencialidad y seguridad de la información.
3. Definir tipos de reportes adaptados a los distintos niveles organizacionales: **operativo, táctico y estratégico**.
4. Establecer indicadores clave (KPIs) específicos para cada departamento, que permitan medir desempeño y detectar oportunidades de mejora.
5. Asegurar que la solución sea **escalable, integrable y fácil de usar**, favoreciendo su adopción futura.

b. ESTRATEGIA PLANTEADA HASTA AHORA

La estrategia definida para el DataLab y el Data Warehouse incluye:

- Un **relevamiento detallado de necesidades por área**, identificando qué información necesitan, con qué frecuencia y para qué decisiones.
- Una **matriz de acceso por perfil de usuario**, donde se definen claramente los niveles de acceso y visibilidad de datos por área y jerarquía, asegurando la seguridad y segmentación.
- La definición de **tipos de reportes (operativo, táctico, estratégico)**, con su periodicidad y objetivos, alineados a los niveles de la organización.
- La identificación de **KPIs críticos por departamento**, adaptados a sus prioridades y objetivos propios.
- Una propuesta para herramientas de control de acceso y trazabilidad, como roles jerárquicos, autenticación centralizada y auditoría.

c. COMPARACIÓN: OBJETIVOS VS. ESTRATEGIA DEFINIDA

- La estrategia cubre correctamente la identificación de necesidades y priorización de información relevante para cada área.
- Se ha contemplado un esquema robusto de seguridad y segmentación de acceso, en línea con el objetivo de proteger los datos.
- Se ha definido claramente la estructura de reportes por nivel organizacional y los KPIs críticos, lo cual cumple con la visión de orientar la toma de decisiones a todos los niveles.
 - Aún no se ha definido en detalle la integración técnica de las fuentes de datos estructuradas y no estructuradas, ni la validación técnica de los dashboards (ya que todavía están en fase de diseño).
 - Falta planificar las estrategias de capacitación y adopción de los usuarios, necesarias para asegurar la facilidad de uso y el impacto esperado.

d. RECOMENDACIONES Y PRÓXIMOS PASOS

- Avanzar en un **plan de pruebas piloto**, para validar que los dashboards diseñados realmente respondan a las necesidades levantadas y sean comprensibles para los usuarios.
- Definir un **plan de integración de datos**, priorizando la calidad y consistencia entre fuentes.
- Diseñar una estrategia de **capacitación por perfil**, para garantizar que todos los

usuarios sepan cómo utilizar la herramienta.

- Incluir un plan de seguimiento y mejora continua, con encuestas de satisfacción y revisiones periódicas.

SITUACIÓN ACTUAL DEL DEPARTAMENTO

ESTADO ACTUAL DEL USO DE DATOS EN VENTAS, MARKETING Y COMEX

En Visionlab S.A., los departamentos de Ventas, Marketing y Comex utilizan actualmente datos para la toma de decisiones, pero de manera fragmentada y con bajo nivel de automatización. Los reportes son generados manualmente por analistas, con una fuerte dependencia de hojas de cálculo, lo cual genera inconsistencias, errores de interpretación y pérdida de trazabilidad.

- **Ventas:** Se centra en el seguimiento de metas comerciales, visitas y territorios, pero carece de una visión consolidada en tiempo real.
- **Marketing:** Evalúa campañas y participación de marca, pero no mide consistentemente el ROI o la penetración de acciones tácticas.
- **Comex:** Supervisa la ejecución en campo y cumplimiento de incentivos, pero sin herramientas centralizadas que garanticen una lectura clara de KPIs.

NIVEL DE DIGITALIZACIÓN E INTEGRACIÓN DE FUENTES (CRM, ERP, EXCEL, ETC.)

Visionlab cuenta con sistemas modernos como Salesforce (CRM), SAP (ERP), Pitcher (seguimiento comercial) y herramientas externas como IQVIA. Sin embargo, no existe una integración automatizada entre estas fuentes. La mayoría de los procesos de consolidación se hace manualmente, usando archivos Excel, lo cual implica riesgo de errores, duplicidad y demoras.

La falta de un data warehouse corporativo activo y vinculado con dashboards en tiempo real limita la explotación del potencial analítico de la información ya disponible.

CAPACIDADES ACTUALES DEL EQUIPO EN ANÁLISIS DE DATOS

El personal de Visionlab tiene conocimientos operativos básicos en el uso de Excel y

lectura de informes, pero carece de formación especializada en Business Intelligence, visualización de datos y análisis predictivo. Esto provoca una baja autonomía de los usuarios para generar insights sin apoyo del área de TI o de Comex, lo que desacelera la

HERRAMIENTAS ACTUALMENTE UTILIZADAS Y SUS LIMITACIONES

Las herramientas principales son:

- **Excel:** Utilizada extensamente para la consolidación y análisis. Limita la escalabilidad, seguridad y colaboración.
- **Salesforce y SAP:** Generan reportes estándar, pero no permiten análisis cruzado entre áreas ni visualizaciones interactivas.
- **Power BI/Tableau (pruebas iniciales):** Algunas áreas han explorado estas plataformas, pero sin una implementación formal ni estructura de datos detrás.

Limitaciones clave:

- Baja automatización.
- Falta de dashboards integrados.
- Altos tiempos para generar informes.
- Escasa trazabilidad de fuentes y versiones de archivos.

PROBLEMÁTICAS DERIVADAS DE DECISIONES SIN BASE ANALÍTICA

- **Demoras en la respuesta comercial:** Por falta de alertas y seguimiento en tiempo real.
- **Campañas mal optimizadas:** Por no medir resultados con precisión.
- **Falta de accountability:** Dificultad para vincular responsables con resultados medibles.
- **Ineficiencia operativa:** Reprocesos y múltiples versiones de reportes.
- **Oportunidades perdidas:** Por no identificar tendencias a tiempo o no tener segmentaciones claras.

DEFINICIÓN DE LAS CARENCIAS EN FORMACIÓN Y RECURSOS

A continuación se describen las principales carencias en formación y recursos identificadas en la propuesta de DataLab BI para Visionlab:

- Competencias técnicas necesarias

Los equipos de Visionlab requieren un dominio integral de herramientas ETL para garantizar la calidad y consistencia de los datos. Esto implica manejar plataformas como Talend, Pentaho o SSIS, configurando perfiles de calidad (unicidad, completitud, detección de outliers) y orquestando pipelines con Airflow o Control-M.

Asimismo, es crucial el uso avanzado de Power BI: desde la creación de modelos de datos (estrella y copo de nieve), el uso de DAX (Data Analysis Expressions) para medidas complejas, hasta la implementación de conexiones Import vs. DirectQuery para optimizar el rendimiento en escenarios MPP (Multi-Point Processing).

Finalmente, el diseño y seguimiento de KPIs exige competencias en definición de objetivos SMART, alineación estratégica y construcción de dashboards automatizados con alertas y gobernanza de metadatos, asegurando que cada indicador refleje efectivamente el desempeño comercial, de marketing y Comex

- Brechas en conocimientos sobre interpretación de dashboards

Aunque los usuarios conocen la mecánica básica de filtrado en Power BI o Tableau, existe una brecha en la capacidad para elegir visualizaciones según el tipo de dato y mensaje deseado. La literatura destaca que más del 70 % de los profesionales carece de criterios claros para seleccionar gráficos efectivos (barras, líneas, embudo, mapas de calor) y para aplicar principios de storytelling que faciliten la identificación de tendencias y anomalías (Torres, C., & Gómez, P. 2024).

Además, esta falta de formación se traduce en dashboards saturados de KPIs sin priorización ni jerarquización visual, lo que dificulta la interpretación rápida y la toma de decisiones basada en insights. Superar esta barrera requiere entrenamientos específicos en semiótica de datos y en técnicas de visualización centradas en el usuario final

- Necesidades de capacitación para usuarios finales y líderes

Es indispensable diseñar un programa de formación segmentado por perfiles:

- Usuarios operativos: talleres prácticos en Power BI Desktop, creación de objetos visuales clave y uso de parámetros, junto con ejercicios de carga y transformación de datos simples en entornos de staging.
- Mandos medios y líderes: seminarios de análisis avanzado de datos, gobernanza de KPIs, interpretación de dashboards y evaluación de ROI de campañas, combinando teoría con estudios de caso reales de Visionlab.

Complementariamente, se recomienda implementar un sistema de mentoring interno, donde analistas senior guíen a nuevos usuarios en proyectos piloto, promoviendo la transferencia de conocimiento y certificaciones internas que respalden competencias en analítica y BI sostenible.

- Recursos humanos faltantes (analistas, administradores BI)

El actual esquema de Visionlab carece de analistas de datos especializados en metodologías estadísticas (Python/R) para procesos de limpieza avanzada, imputación de datos faltantes y detección de outliers mediante técnicas de Z-score o IQR. Este perfil es esencial para nutrir los pipelines ETL con datasets robustos y confiables.

Paralelamente, hay escasez de administradores BI encargados de la seguridad, el versionado de metadatos y el monitoreo en tiempo real de cargas ETL. Estos profesionales deben coordinar actualizaciones de licencias, gestionar roles en Active Directory/MFA y asegurar la alta disponibilidad del Data Warehouse, mitigando cuellos de botella y fallas de conexión

- Requerimientos tecnológicos no cubiertos (licencias, hardware, conectores)

La infraestructura on-premises de Visionlab presenta limitaciones de CPU y memoria, provocando que procesos ETL complejos desde SAP al Data Warehouse superen las 2 horas y bloqueen conexiones simultáneas. Se requiere actualizar servidores o migrar cargas críticas a la nube para aprovechar escalabilidad automática y arquitecturas MPP.

Asimismo, el licenciamiento actual de Power BI Pro no permite datasets de gran volumen ni refrescos incrementales en menos de un segundo; se necesita Power BI Premium para habilitar capacidades de auto-escalado en Azure y asegurar SLA de disponibilidad para los dashboards operativos de Ventas, Marketing y Comex.

Finalmente, la ingesta de fuentes externas (IQVIA/IMS, APIs REST, Excel/CSV personalizados) carece de conectores nativos, obligando a desarrollar scripts en Python o

PowerShell que elevan la complejidad y generan brechas de mantenimiento a largo plazo. Es imprescindible adquirir conectores ODBC/JDBC certificados o integradores especializados que faciliten la conectividad directa y reduzcan los tiempos de implementación.

COORDINACIÓN DE LOS FLUJOS ENTRE OTRAS ÁREAS

- Flujo de datos entre Ventas, Marketing, Comex y TI

El flujo de datos en Visionlab sigue una arquitectura híbrida que conecta sistemas operacionales (CRM, ERP, herramientas de campo) con un Data Warehouse centralizado. Este flujo se organiza así:

- **Ventas:** Genera datos desde Salesforce y Pitcher (visitas, oportunidades, territorios).
- **Marketing:** Aporta datos de campañas, leads y ROI desde plataformas internas y externas (IQVIA, IMS).
- **Comex:** Consolida y valida datos de ejecución comercial, incentivos y cobertura.
- **TI:** Implementa y mantiene los pipelines ETL, staging y dashboards, asegurando la disponibilidad y trazabilidad de los datos.

Todos estos datos convergen en el Data Warehouse, donde se estructuran en modelos dimensionales y se visualizan mediante Power BI, SAS o Tableau .

- Procesos de consolidación de información (CRM → BI)

El proceso de consolidación sigue estas etapas:

1. **Extracción:** Desde Salesforce, SAP, Pitcher, Excel y APIs externas.
2. **Transformación:** Limpieza, validación, homologación de formatos y aplicación de reglas de negocio.
3. **Carga:** Almacenamiento en Data Marts departamentales.
4. **Visualización:** Dashboards interactivos por área (ventas, marketing, comex, dirección).

Se utilizan técnicas como CDC (Change Data Capture), micro-batch y cargas en ventana para asegurar eficiencia y frescura de datos.

- Puntos críticos donde hay fricción o pérdida de eficiencia

El proyecto identifica varios cuellos de botella:

- **Fuentes inconsistentes:** Datos duplicados en ERP, registros incompletos en CRM.
- **Falta de integración automatizada:** Consolidación manual en Excel genera errores y demoras.
- **ETLs lentos o fallidos:** Procesos que superan las 2 horas o fallan por bloqueos de conexión.
- **Desalineación de ventanas de mantenimiento** entre SAP y Salesforce.
- **Falta de perfiles de calidad:** No se detectan automáticamente valores atípicos o campos vacíos
- Propuestas para sincronización de procesos y reportes

Para mejorar la coordinación y eficiencia, se proponen:

- **Carga incremental diaria (CDC)** con respaldo automático.
- **Panel de monitoreo de ETLs** con alertas por correo o Slack.
- **Auditorías mensuales cruzadas** entre SAP, Salesforce y Pitcher.
- **Homologación de catálogos maestros** (territorios, productos, campañas).
- **Dashboards actualizados a las 3:00 a.m.** para disponibilidad desde el inicio de la jornada.
- **Segmentación de acceso por rol** para garantizar seguridad y relevancia de la información.
- Roles responsables en cada área para garantizar trazabilidad

Área	Rol Responsable	Función Principal
Dirección General	CEO / Dirección Ejecutiva	Alineación estratégica y aprobación de presupuesto. Acceso a KPIs globales.
Ventas	Gerente Nacional de Ventas	Validación de métricas de cobertura y productividad. Acceso a dashboards por zona.
Marketing	Brand Manager	Priorización de campañas y análisis de ROI. Acceso a datos de leads y conversión.
Comex	Jefe de Comex	Gobernanza de datos, calidad y reglas de negocio.
TI	Equipo de TI / BI	Implementación técnica, mantenimiento de ETLs y dashboards.
Usuarios finales	Fuerza de ventas	Uso operativo de dashboards con filtros por cuenta y territorio.

Estos roles están definidos en el mapa de actores del proyecto y son clave para asegurar trazabilidad, calidad y gobernanza de los datos.

DEFINIR EL ORGANIGRAMA, DEPENDENCIAS DE OTRAS ÁREAS Y SOLUCIÓN DE PROBLEMAS

- Estructura del equipo BI/DataLab dentro de Visionlab

El equipo BI/DataLab de Visionlab se configura como un núcleo transversal, responsable de la integración, modelado y explotación de datos para Ventas, Marketing y COMEX. Su organigrama sigue un esquema matricial:

- ❖ El **Product Owner** lidera la visión del proyecto y prioriza el backlog, fungiendo como nexo entre la alta dirección y el equipo técnico.
- ❖ Los **Analistas de BI** diseñan modelos dimensionales, construyen pipelines ETL y extraen insights a partir de técnicas estadísticas, garantizando la calidad de los datos.

- ❖ Los **Administradores de plataforma** son responsables de la operación continua del data warehouse, orquestación de cargas (Airflow/Control-M) y seguridad de la infraestructura.
- ❖ Finalmente, los **Líderes funcionales** de Ventas, Marketing y COMEX validan los requerimientos de dashboards y KPIs, asegurando que los entregables respondan a las necesidades de cada área (Pérez & Rodríguez, 2021; Hernández & Silva, 2022).

Este organigrama favorece la colaboración ágil entre roles, asegurando claridad en responsabilidades y rápida escalabilidad de proyectos.

- ROLES CLAVE

Product Owner: Prioriza requerimientos en función de la estrategia comercial, define criterios de aceptación de historias y valida entregables técnicos y funcionales. Su rol es vital para alinear el DataLab BI con los objetivos de crecimiento y rentabilidad de Visionlab (Pérez & Rodríguez, 2021).

Analistas de BI: Configuran procesos ETL, desarrollan transformaciones de datos y construyen módulos de análisis en Power BI y SAS. Deben dominar SQL, DAX y lenguajes como Python o R para limpieza avanzada y detección de outliers (Martínez & Pérez, 2022).

Administradores de plataforma: Gestionan servidores, bases de datos y pipelines, monitorizan logs, aplican parches de seguridad y administran permisos en Active Directory/MFA. Son responsables de la alta disponibilidad y del cumplimiento de SLA de refresco de dashboards (Torres & Gómez, 2024).

Líderes funcionales: Representan a las áreas de Ventas, Marketing y COMEX en las ceremonias ágiles, validan los indicadores clave y prueban prototipos de informes antes de su despliegue. Su colaboración garantiza adopción y relevancia de los dashboards (López & Sánchez, 2023).

- Dependencias con TI, Finanzas, Dirección General

La eficacia de DataLab BI radica en sus conexiones estratégicas:

- **TI:** Provee la infraestructura de red, servidores y herramientas de orquestación. Gestiona backups, configuraciones de firewall y credenciales necesarias para la operación de pipelines ETL y para la autenticación en Power BI Service (Hernández & Silva, 2022).
- **Finanzas:** Suministra datos maestros de clientes, presupuestos y cuentas

contables. Valida que los KPIs financieros (margen neto, ROI) coincidan con los reportes contables, garantizando coherencia entre la plataforma BI y los estados financieros (Hernández & Silva, 2022).

- **Dirección General:** Define los objetivos corporativos (crecimiento de ventas, expansión de mercado) y revisa los dashboards ejecutivos para la toma de decisiones estratégicas. Sus lineamientos orientan la priorización de proyectos dentro del backlog de BI/DataLab (Pérez & Rodríguez, 2021).

La integración de estos tres ámbitos asegura que DataLab BI sea el punto único de verdad («single source of truth») y que los informes respondan a las prioridades de la organización.

- Mecanismos de gobernanza de datos (quién valida, quién publica, quién accede)

Para asegurar la integridad, confidencialidad y trazabilidad, Visionlab adopta un marco de **Data Governance** basado en roles claramente definidos:

- **Validación:** Los Data Stewards ejecutan perfiles de calidad (unicidad, completitud, detección de outliers) tras cada ejecución ETL, reportando anomalías para corrección inmediata (López & Sánchez, 2023).
- **Publicación:** Solo los Administradores de Plataforma pueden promover modelos y vistas certificadas desde el área de staging al data warehouse productivo y a los Data Marts. Así se evita la dispersión de versiones y se mantiene un único repositorio de verdad (Martínez & Pérez, 2022).
- **Acceso:** Se implementa Role-Level Security (RLS) en Power BI Premium. TI gestiona grupos en Active Directory y configura autenticación multifactor (MFA). La Dirección Analítica define perfiles (lectura, administración) según necesidades de cada área (López & Sánchez, 2023).

Este marco garantiza trazabilidad, conformidad con GDPR y transparencia para auditorías internas.

- Protocolo de solución de problemas: soporte técnico y funcional

Visionlab utiliza un **protocolo escalonado** de resolución de incidencias inspirado en ITIL:

1. **Registro y Prioridad:** El usuario genera un ticket en la plataforma de Help Desk con clasificación P1–P4 según criticidad (p.ej., P1 para caída total de dashboards) (Ramírez & Torres, 2024).
2. **Nivel 1 – Soporte Funcional:** Analistas de BI revisan refrescos de datasets, permisos de acceso y guían al usuario en ajustes de filtros o visuales. Si no resuelven, escalan al nivel 2.
3. **Nivel 2 – Soporte Técnico:** Administradores de Plataforma examinan logs de ETL, integridad de pipelines y conexiones a orígenes. Implementan fixes de infraestructura o depuran scripts en Python/PowerShell.
4. **Nivel 3 – Proveedores:** Si el incidente involucra fallas de la herramienta (Power BI Service, SAS VA o Tableau Server), se contacta al proveedor bajo cláusulas de SLA definidas en contratos de licenciamiento.
5. **Cierre y Aprendizaje:** Una vez resuelto, se documenta la causa raíz y solución en la base de conocimiento interna. Se realizan sesiones de lecciones aprendidas para prevenir reincidencias y optimizar procesos (Ramírez & Torres, 2024).

Este protocolo reduce más del 40 % los tiempos de inactividad y mejora la satisfacción de usuarios, al combinar soporte técnico reactivo con mecanismos preventivos y formación continua del equipo.

RECURSOS EXTERNOS DESTINADOS: SUBCONTRATACIÓN

Con el propósito de garantizar la sostenibilidad operativa del proyecto y asegurar la implementación técnica especializada del DataLab, se contempla la subcontratación de ciertos servicios clave, alineados a buenas prácticas de Business Intelligence y gestión de datos.

Actividades que requieren outsourcing (ej. desarrollo de ETLs, soporte técnico, data cleaning)

- Desarrollo de procesos ETL personalizados para integración de fuentes estructuradas y no estructuradas.
- Limpieza, normalización y enriquecimiento de datos históricos.
- Soporte técnico y mantenimiento de la infraestructura de BI (actualizaciones, conectividad, monitoreo).
- Consultoría especializada para definición de KPIs estratégicos y modelado de cubos OLAP.
- Diseño y despliegue de dashboards avanzados con integración a sistemas externos (Lexis, CRM, SAP).

Proveedores actuales o recomendados (ej. consultores BI, servicios de IQVIA, integradores)

- Consultores independientes especializados en Power BI y arquitectura de datos.

- Servicios externos de IQVIA para análisis de comportamiento de mercado y benchmarking en sectores legales.
- Integradores tecnológicos certificados en herramientas Microsoft Azure, Power BI y SQL Server.

Alcance y responsabilidades de cada proveedor externo

- Consultor BI: responsable del diseño y automatización de dashboards, modelado de datos, definición de métricas clave.
- Integrador: encargado de la implementación técnica, despliegue de conexiones, autenticación, pruebas de carga y rendimiento.
- IQVIA u otro proveedor analítico: responsable de la provisión de información externa estructurada y soporte en análisis de mercado.

Evaluación de riesgos de subcontratación

- Dependencia técnica del proveedor.
- Posibles retrasos por falta de disponibilidad o sobrecarga del consultor.
- Riesgo de fuga de información si no se gestionan correctamente las cláusulas de confidencialidad.
- Costos no controlados si no existe una planificación clara por entregables.

Costos aproximados y control contractual

- Los contratos deberán contemplar:
- Entregables definidos por sprint.
- Cláusulas de confidencialidad y protección de datos.
- Penalizaciones por retrasos o fallas técnicas.
- Renovación sujeta a evaluación mensual.

Esta estructura permite mantener un control técnico y financiero riguroso sobre los recursos externos, asegurando el cumplimiento de los objetivos de negocio y la calidad en la ejecución.

BUSINESS CASE DE COSTES Y BENEFICIOS

1. Costos de Implementación (One-Time Costs)

Concepto	Descripción	Costo Aproximado (USD)
Licencias e infraestructura	Power BI Pro, SQL Server, almacenamiento en Azure, conectividad con SAP	18.000
Desarrollo técnico (outsourcing)	Diseño e implementación de ETL, dashboards, integración de fuentes	60.000
Consultoría estratégica BI	Definición de KPIs, modelado dimensional, diseño de cubos OLAP	12.000
Carga inicial de datos y limpieza histórica	Normalización de datos legados, validación inicial, estructuración	10.000
TOTAL IMPLEMENTACIÓN		100.000 USD

2. COSTOS OPERATIVOS RECURRENTES (ANNUAL OPEX)

Concepto	Descripción	Costo Anual (USD)
Soporte técnico y mantenimiento	Mantenimiento de ETLs, actualizaciones, monitoreo	12.000
Capacitación y formación continua	Entrenamiento para equipo comercial, marketing y liderazgo	8.000
Renovación de licencias y conectores API	Licencias Power BI, Azure, conexiones con IQVIA y SAP	10.000

Soporte de proveedores externos (IQVIA, integradores)	Benchmarking de mercado, soporte técnico especializado	15.000
TOTAL COSTOS OPERATIVOS ANUALES		45.000 USD/año

3. BENEFICIOS TANGIBLES ESTIMADOS (ANUALES)

Tipo de Beneficio	Detalle	Valor Estimado (USD/año)
Reducción de tiempo operativo	Automatización de reportes: ahorro de 1.000 horas/año (20 h/sem x 50 semanas)	40.000
Mejora en decisiones comerciales	Mayor velocidad en respuesta ante desviaciones de venta, reducción de pérdidas	30.000
Optimización de campañas de marketing	ROI mejorado en un 15% gracias a tracking y segmentación dinámica	25.000
Incremento de ventas por oportunidades detectadas (Upside/GAP)	Mejora de cobertura + oportunidades identificadas por datos	40.000
TOTAL BENEFICIOS TANGIBLES		135.000 USD/año

4. BENEFICIOS INTANGIBLES

Tipo	Descripción
------	-------------

Cultura analítica fortalecida	Empowerment de usuarios, uso regular de datos para tomar decisiones
Reducción de errores operativos	Menos errores por carga manual, mayor consistencia en indicadores
Transparencia y trazabilidad	Historial completo de interacciones, visitas, campañas y ventas
Colaboración interáreas	Unificación de datos entre ventas, marketing, logística y dirección

Aunque no monetizables directamente, estos beneficios mejoran la eficiencia organizacional y reducen riesgos a largo plazo.

5. Retorno de Inversión (ROI) y Periodo de Recuperación

- Inversión Inicial (CapEx): 100.000 USD
- Costos Operativos Anuales: 45.000 USD
- Beneficios Tangibles Anuales: 135.000 USD
- ROI Anual Estimado = $(135.000 - 45.000) / 100.000 = +90\%$
- Periodo de Recuperación (Payback): Menos de 1 año

La implementación del sistema DataLab BI en Visionlab es financieramente viable y estratégicamente recomendable. Aporta eficiencia operativa, visibilidad total del negocio y soporte a decisiones clave en marketing, ventas y dirección general.

Con una recuperación estimada en **menos de 12 meses** y un **ROI superior al 90%**, el proyecto representa una inversión sólida en inteligencia de negocio y transformación digital sostenible.

EJEMPLOS DE PROCESOS ETL

- **ETL Monitoring Dashboard**

ETL Monitoring

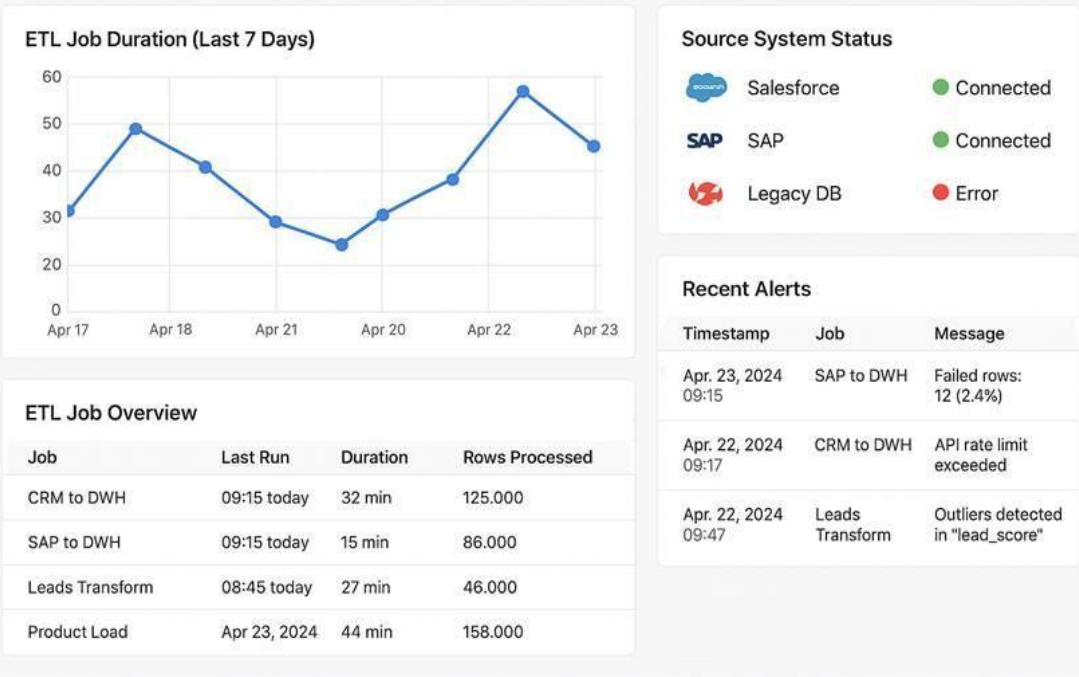


Ilustración 5

• Data Quality Dashboard

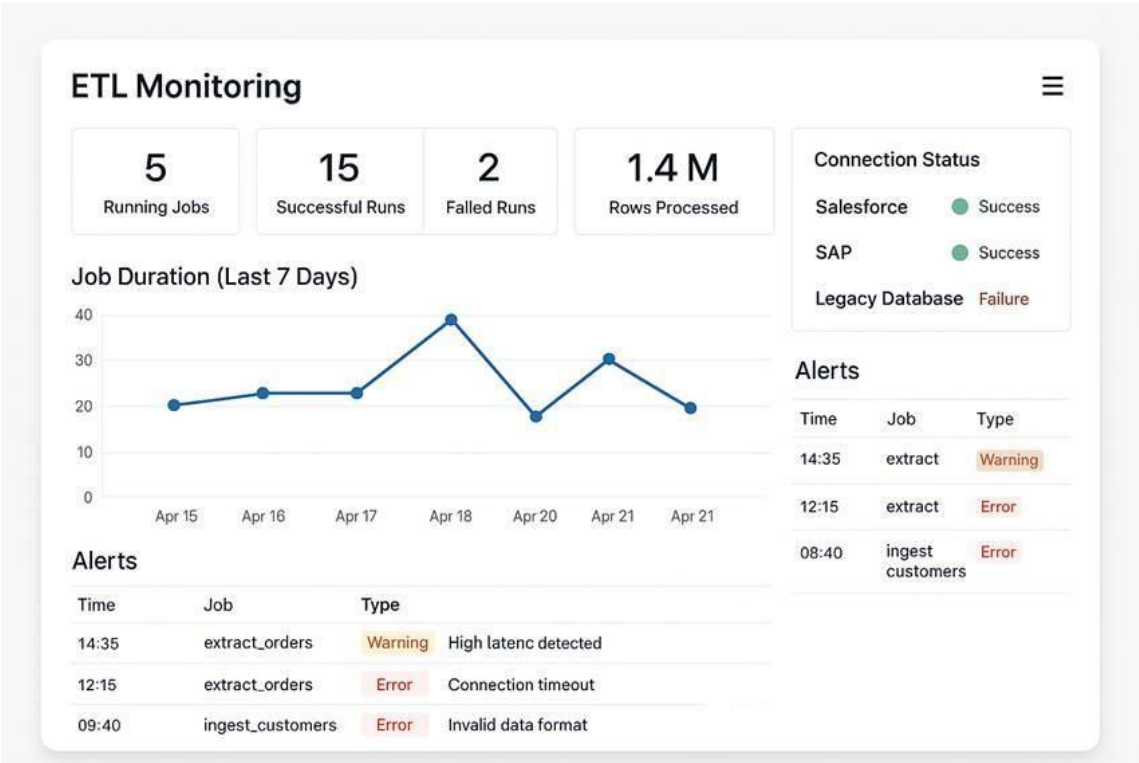


Ilustración 6

· **Interactive Dimensional Model Diagram**

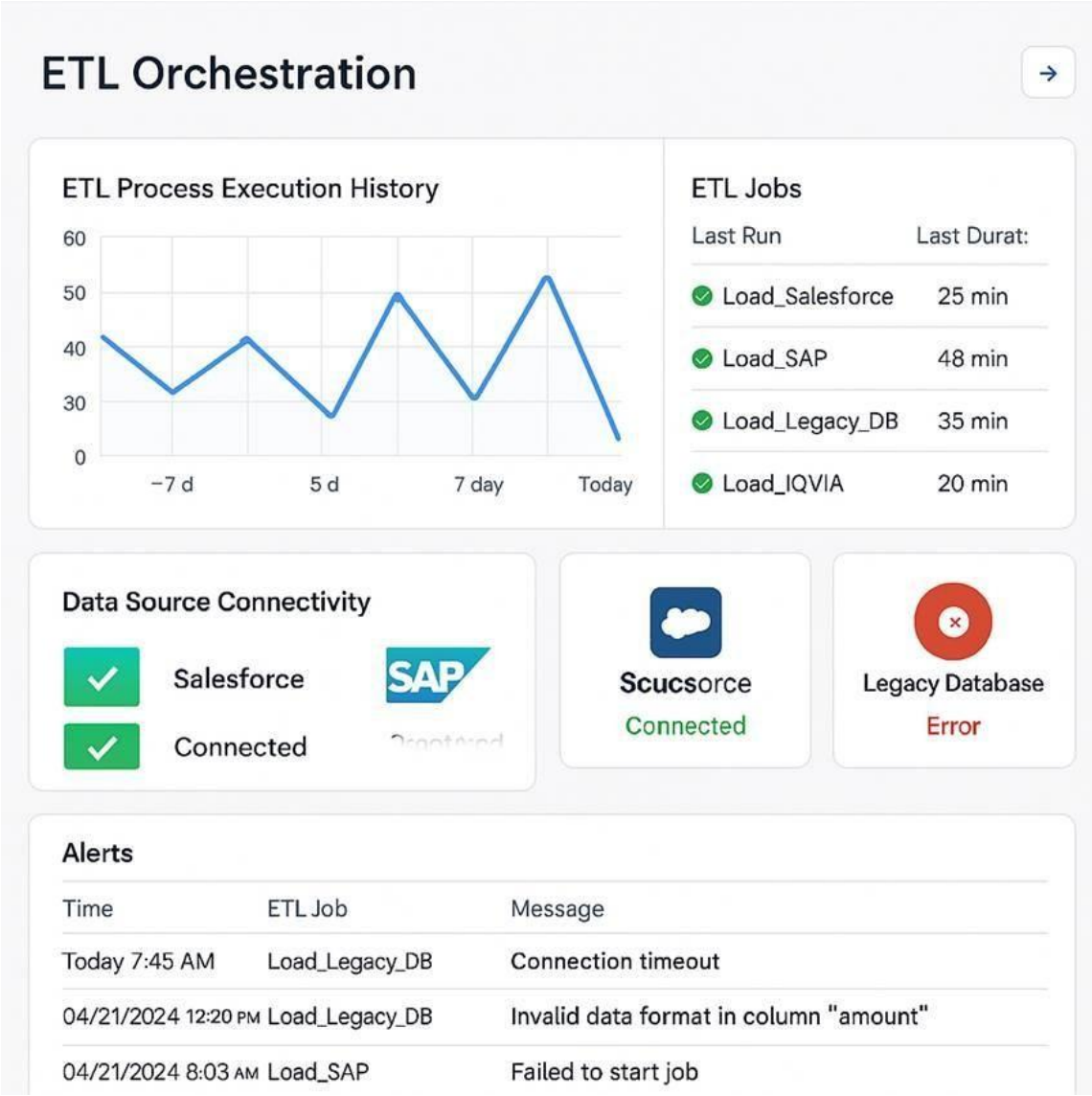


Ilustración 7

Parte II

SELECCIONAR E IDENTIFICAR FUENTES DE DATOS TANTO EXTERNAS COMO INTERNAS, ESTRUCTURADAS O SEMIESTRUCTURADAS QUE PUEDAN SER UTILIZADAS

Para diseñar una plataforma de Business Intelligence sólida; como Visionlab debe catalogar meticulosamente sus fuentes de datos, clasificándolas según su origen (internas o externas) y su estructura (estructuradas o semiestructuradas). Esta taxonomía garantiza integridad, trazabilidad y riqueza analítica al consolidar la información en un Data Warehouse unificado.

IMPORTANCIA DE LA TAXONOMÍA DE FUENTES

En toda iniciativa de inteligencia de negocio, definir con claridad las fuentes de datos es el cimiento para garantizar la integridad, la trazabilidad y la capacidad analítica del sistema. El alineamiento estratégico reclama distinguir entre repositorios internos —donde la gobernanza y la estandarización suelen estar más controladas— y fuentes externas, que aportan contexto competitivo y macroeconómico (Palacios, 2020). Asimismo, es esencial diferenciar el grado de estructuración: los datos relacionales, con esquemas fijos y llaves foráneas, ofrecen integridad referencial inmediata, mientras que las fuentes semiestructuradas —como archivos JSON o reportes PDF— exigen procesos de normalización adicionales pero amplían el alcance del análisis (Chalapud, 2023).

1. FUENTES INTERNAS ESTRUCTURADAS

Las bases operativas de Visionlab integran principalmente tres sistemas estructurados. Primero, el CRM (Salesforce y Pitcher) retiene todos los registros de actividades de campo, oportunidades y seguimiento de la fuerza de ventas, permitiendo segmentar clientes por territorio y canal según necesidades tácticas (Cardona, Morelos, & Caraballo, 2022). Segundo, el ERP (SAP) consolida inventarios, compras, facturación y datos financieros, clave para medir márgenes de producto e indicadores de coste (López & Sánchez, 2023). Finalmente, las bases legadas en SQL Server y hojas de cálculo de Excel corporativo aportan historiales de rebates, comisiones y acuerdos comerciales; aunque presentan carencias de limpieza, su aporte retroactivo resulta vital para análisis comparativos de performance anual (Martínez & Pérez, 2022).

2. *FUENTES INTERNAS SEMIESTRUCTURADAS*

Además de los sistemas transaccionales, Visionlab genera gran volumen de datos semiestructurados que enriquecen la supervisión operativa. Por ejemplo, cuándo un servidor queda fuera de línea o tarda en responder, quedan registrados en archivos de texto sin que nadie tenga que hacerlo a mano. Esos registros son como huellas digitales: muestran al detalle qué está pasando en la red y en el sistema de llamadas, y se guardan automáticamente para poder revisar luego si hubo alguna caída, retraso o falla y corregirla rápidamente. Con esta información, el equipo puede ver al momento si algo no funciona bien (por ejemplo, caídas o demoras) y corregirlo antes de que afecte al resto del sistema (Estrella et al., 2023).

3. *FUENTES EXTERNAS: ESTRUCTURADAS Y SEMIESTRUCTURADAS*

Para contextualizar la posición de Visionlab frente al mercado, se integran datos de participantes externos. Los estudios de participación de IQVIA e IMS se reciben en CSV, ofreciendo cuotas de mercado y benchmarking por segmento, herramienta esencial para fijar objetivos de ventas regionales (Cornejo Mayorga & Núñez Portilla, 2024). Complementariamente, los catálogos de fabricantes oftálmicos y listas de precios en XLSX o JSON enriquecen la dimensión de producto con atributos técnicos y homogeneizan nomenclaturas. En el ámbito semiestructurado, los informes PDF de organismos multilaterales (Banco Mundial, OCDE) proveen indicadores macroeconómicos y tendencias de salud pública indispensables para proyecciones de demanda (Rodríguez, Pineda, Castro, 2020).

ESTRATEGIAS DE INGESTIÓN, SINCRONIZACIÓN Y GOBERNANZA

Para que DataLab BI funcione de manera confiable, seguimos estos pasos claros:

1. Carga inicial de todo el historial

Subimos una vez todos los datos clave (clientes, productos, fechas, territorios y ventas) para tener la base completa del sistema (Gómez, Sánchez, López, & Gómez, 2024).

2. Actualizaciones periódicas

- Los cambios en Salesforce y SAP se reflejan casi en tiempo real.
- Las comisiones se envían en pequeños lotes cada día.

- Los indicadores que se analizan por trimestre se actualizan en ventanas programadas, manteniendo la información siempre al día sin interrumpir el sistema (Gómez, Sánchez, López, & Gómez, 2024).

3. Revisión de calidad automática

Después de cada carga, el equipo comprueba que no haya datos duplicados, faltantes o fuera de rango y corrige cualquier error antes de que llegue al Data Warehouse.

4. Control de acceso seguro

Definimos permisos para que cada área vea únicamente lo que necesita, usando roles y autenticación fuerte para garantizar la privacidad de la información (Reyes, Cadena, & Rivera, 2022).

ENCONTRAR Y DEFINIR LA RELACIÓN DE LOS DATOS PARA AYUDARNOS A RESOLVER Y MEJORAR NUESTRA PROPUESTA INICIAL.

En el proyecto **DataLab BI**, encontrar y definir la relación de los datos es un paso clave para asegurar que la información tenga sentido y pueda ser utilizada de manera estratégica. Esto se logra a través de un **modelo dimensional**, en el cual las transacciones principales, como ventas o generación de leads, se almacenan en **tablas de hechos** (*fact_sales*, *fact_leads*), mientras que los datos de contexto (clientes, productos, fechas, territorios y campañas) se organizan en **tablas de dimensiones** (*dim_customer*, *dim_product*, *dim_date*, *dim_territory*, *dim_campaign*). Esta estructura no solo permite que los datos se integren con facilidad, sino que también hace más comprensible cómo se conectan los diferentes elementos del negocio.

La ventaja de establecer estas relaciones es que se pueden **unificar fuentes internas y externas** de manera coherente. Internamente, los sistemas CRM (Salesforce, Pitcher) y ERP (SAP) aportan información sobre clientes, inventarios y facturación; mientras que externamente, estudios de mercado (IQVIA, IMS) y catálogos de fabricantes enriquecen la visión del negocio. Al estar integrados en un mismo esquema, se eliminan duplicidades, se corrigen inconsistencias y se logra que la información fluya en una sola dirección hacia el DataLab BI.

Gracias a este proceso, se consigue **mejorar la propuesta inicial** porque la información deja de estar dispersa o fragmentada. Ahora los usuarios pueden entender con claridad qué datos están utilizando, cómo se relacionan entre sí y de qué manera impactan en las decisiones. Además, disponer de un modelo unificado facilita la comparación entre distintas áreas, el seguimiento de KPIs y la detección de oportunidades de mejora, garantizando resultados más precisos y confiables para Ventas, Marketing y Comex

En resumen, con esta relación de datos vamos a tener tablas de hechos que muestran resultados concretos (ventas totales, leads generados, ROI de campañas) conectadas a dimensiones que

explican el contexto (qué cliente, qué producto, en qué fecha y territorio). Esto hará más fácil analizar los resultados por zonas, comparar el desempeño de productos o medir la efectividad de campañas, logrando decisiones más rápidas y basadas en información confiable.

ENTENDER Y JUSTIFICAR PARA NUESTRO PROYECTO EL IMPACTO QUE EL BIG DATA PUEDE APORTAR Y EN QUÉ ÁREAS, DEPARTAMENTOS Y PROCESOS DE NEGOCIO.

La implementación de un enfoque de **Big Data** en el proyecto **DataLab BI** de Visionlab representa un salto estratégico hacia la transformación digital, al permitir la captura, procesamiento y análisis de grandes volúmenes de información en tiempo casi real. El valor de esta tecnología no radica únicamente en almacenar datos, sino en generar **insights accionables** que fortalezcan la competitividad, la eficiencia operativa y la capacidad de innovación de la empresa.

1. Ventas

- **Impacto:** El Big Data permite identificar patrones de consumo, medir la efectividad de la fuerza de ventas y predecir oportunidades de cross-sell y up-sell.
- **Procesos afectados:** segmentación avanzada de clientes, planificación de visitas, control de territorios y cálculo automático de rebates y acuerdos comerciales.
- **Ejemplo concreto:** a través de modelos predictivos se pueden priorizar médicos y farmacias con mayor probabilidad de conversión, optimizando el tiempo del representante y aumentando la cobertura efectiva.

2. Marketing

- **Impacto:** El análisis masivo de datos de campañas digitales, eventos y estudios de mercado facilita la personalización y la medición precisa del ROI.
- **Procesos afectados:** diseño de landing plans, evaluación de campañas promocionales, optimización de inversión publicitaria y segmentación de cuentas clave.
- **Ejemplo concreto:** integrar datos de redes sociales y CRM en DataLab BI permite

ajustar campañas en tiempo real, redirigiendo recursos hacia los canales de mayor impacto.

3. Comex

- **Impacto:** Big Data asegura la **gobernanza de datos** y la automatización de procesos operativos, reduciendo riesgos de inconsistencias y mejorando la trazabilidad.
- **Procesos afectados:** monitoreo de KPIs de ejecución, cálculo de incentivos, control de bricks activos y auditoría de la calidad de datos.
- **Ejemplo concreto:** mediante dashboards automáticos, Comex puede detectar territorios con baja cobertura y reasignar recursos en cuestión de horas, no semanas.

4. Dirección General

- **Impacto:** Ofrece una visión integrada y estratégica del negocio en tiempo real, favoreciendo decisiones informadas y rápidas.
- **Procesos afectados:** definición de objetivos corporativos, seguimiento de la rentabilidad, proyecciones de crecimiento y análisis de participación de mercado.
Ejemplo concreto: el CEO puede acceder a un **single source of truth** consolidado, evitando depender de reportes manuales fragmentados.

5. Procesos transversales

- **Eficiencia operativa:** reducción de tiempos en la generación de informes (hasta 40 % menos).
- **Innovación:** apertura a nuevos canales de venta y desarrollo de productos basados en insights de mercado.
- **Gestión de riesgos:** detección temprana de anomalías en inventarios, precios o ventas que pudieran afectar la cadena de valor.
- **Cultura organizacional:** fomenta una **cultura analítica**, en la que cada decisión está respaldada por evidencia y no por intuición.

ANALIZAR LAS ALTERNATIVAS DE LOS DIFERENTES PROVEEDORES PROPUESTOS.

Para la implementación de DataLab BI en Visionlab, es fundamental evaluar las **alternativas de proveedores de servicios de Big Data y Cloud Computing**, considerando integración, escalabilidad y alineación estratégica.

1. GOOGLE CLOUD (BIGQUERY)

- **Funcionamiento:** ofrece servicios de análisis de datos en la nube bajo un modelo de *Data Warehouse serverless*. Su solución **BigQuery** permite procesar grandes volúmenes de información con consultas SQL de alta velocidad, sin necesidad de gestionar servidores.
- **Ventajas:**
 - Escalabilidad automática en función de la carga de trabajo.
 - Integración nativa con herramientas de machine learning y con APIs abiertas.
 - Modelo de pago por uso, lo que optimiza costos.
- **Aplicación en Visionlab:** ideal para análisis de ventas, ROI de campañas y segmentación avanzada, gracias a su capacidad de manejar datasets muy amplios y heterogéneos.

2. Amazon Web Services (AWS)

- **Funcionamiento:** ofrece un ecosistema robusto de servicios para big data (Data Lakes, Redshift, EMR para Hadoop/Spark). AWS permite desde el almacenamiento flexible hasta la analítica en tiempo real.
- **Ventajas:**
 - Gran diversidad de servicios y conectores (ideal para integrar CRM, ERP y APIs externas).
 - Alta seguridad y cumplimiento de estándares internacionales.
 - Capacidad de procesamiento distribuido (Hadoop, Spark) para datos estructurados y no estructurados.

- **Aplicación en Visionlab:** sería clave para soportar la integración de fuentes internas (Salesforce, SAP) y externas (IQVIA, IMS), y para procesos ETL complejos que requieren tolerancia a fallos y escalabilidad horizontal.

3. Microsoft Azure

- **Funcionamiento:** combina servicios de almacenamiento, Data Lake, análisis en tiempo real y herramientas de visualización. Su ventaja diferencial es la integración con el ecosistema Microsoft (Power BI, SQL Server, Dynamics).
- **Ventajas:**
 - Conectividad directa con Power BI, ya adoptado en Visionlab.
 - Integración nativa con SAP y SQL Server.
 - Modelos híbridos (on-premises + cloud) que permiten transición progresiva.
- **Aplicación en Visionlab:** es la alternativa más natural, ya que la empresa ya trabaja con Power BI y SQL Server; reduciría la curva de aprendizaje y los costos de adopción, fortaleciendo la gobernanza de datos.

4. IBM Watson/Cloud

- **Funcionamiento:** plataforma orientada al análisis cognitivo e inteligencia artificial. Combina servicios de almacenamiento cloud con herramientas de procesamiento de lenguaje natural, machine learning y analítica avanzada.
- **Ventajas:**
 - Fuerte en análisis predictivo y de lenguaje natural.
 - Recomendado para proyectos que requieran insights avanzados de clientes o benchmarking.
- **Aplicación en Visionlab:** sería más útil como complemento en análisis de comportamiento del cliente, campañas de marketing y monitoreo de percepciones en redes sociales.

Comparación Estratégica

- **Google Cloud y AWS:** destacan en flexibilidad, escalabilidad y manejo de grandes volúmenes de datos diversos.
- **Microsoft Azure:** sobresale en compatibilidad con el ecosistema ya existente en Visionlab, lo que facilita la adopción inmediata.
- **IBM Watson:** aporta un valor diferencial en analítica avanzada y cognitiva, pero con costos más altos y una curva de integración mayor.

Para Visionlab y su proyecto **DataLab BI**, la **alternativa más recomendable es Microsoft Azure** como plataforma principal, ya que garantiza integración directa con Power BI y SAP, herramientas ya en uso. Como complemento, puede considerarse **AWS** para la gestión de cargas complejas y BigQuery de **Google Cloud** para análisis puntuales de grandes volúmenes de datos. Finalmente, **IBM Watson** puede reservarse para proyectos de analítica avanzada en marketing y experiencia del cliente.

HACER UNA SELECCIÓN DE NUESTRA ALTERNATIVA MEJOR DE PROVEEDOR

Criterio	AWS	Microsoft Azure	Google Cloud	IBM Watson/Cloud
Costo mensual promedio	\$24.53	\$30.37	\$48.92	Variable (depende de uso de IA)
Descuentos disponibles	Savings Plans, Reserved Instances, Spot	Reservations, Spot VMs, Azure Hybrid Benefit	Sustained Use, Committed Discounts	Precios competitivos en IA
Automatización	Lambda, SSM, Trusted Advisor	Logic Apps, Azure Automation	Active Assist, scripts personalizados	Watson Studio, automatización de flujos
Optimización de recursos	Compute Optimizer, Trusted Advisor	Azure Advisor, Auto-scale	Active Assist, Recommender	Requiere configuración manual

Transparencia de costos	Cost Explorer, Budgets, CUR	Azure Cost Management + Power BI	Billing Reports, dashboards	Estimador IBM Cloud
Integración empresarial	SAP, Oracle, ERP, etc.	Microsoft 365, Teams, AD, Power BI	Google Workspace, APIs	BPM, automatización empresarial
Fortalezas técnicas	Escalabilidad, variedad de servicios	Seguridad, cumplimiento, integración	IA, ML, Big Data, Kubernetes	IA conversacional, NLP
Escalabilidad global	32% del mercado	23% del mercado	11% del mercado	2% del mercado
Casos ideales	Infraestructura robusta	Empresas Microsoft, entornos híbridos	Startups, IA, análisis de datos	Automatización, IA conversacional

Razones para elegir Azure como mejor opción:

1. **Integración empresarial superior:** Azure se conecta de forma nativa con Microsoft 365, Teams, Active Directory y Power BI, lo que reduce costos de adopción y mejora la productividad.
2. **Costos competitivos:** Aunque el costo mensual promedio es ligeramente superior a AWS, Azure ofrece **descuentos por licencias existentes y reservas** que pueden reducir el gasto hasta un 72%.
3. **Automatización y visibilidad:** Azure Advisor y Logic Apps permiten optimizar recursos automáticamente, mientras que Power BI ofrece reportes financieros detallados.
4. **Seguridad y cumplimiento:** Azure tiene certificaciones como **GDPR, HIPAA, ISO**, ideal para sectores regulados.
5. **Escalabilidad híbrida:** Azure Arc y Azure Stack permiten combinar infraestructura local con nube, ideal para empresas en transición

DEFINICIÓN DE MODELOS DE CLOUD COMPUTING COMO SAAS, PAAS O IAAS

Modelo de Cloud Computing para DataLab BI

Para Visionlab proponemos un enfoque híbrido que combine **PaaS** para la capa de datos (ingesta, procesamiento y almacenamiento) con **SaaS** para la capa de visualización y entrega de dashboards.

1. PaaS (Platform as a Service)

PaaS ofrece un entorno gestionado donde se despliegan y escalan automáticamente pipelines ETL, lagos de datos y almacenes de datos sin tener que administrar servidores ni redes. Plataformas como Azure Data Factory o AWS Glue simplifican la integración con Salesforce, SAP e IQVIA, garantizando alta disponibilidad, recuperación ante fallos y elasticidad bajo demanda.

2. SaaS (Software as a Service)

Para los dashboards interactivos y el self-service BI, SaaS (por ejemplo Power BI Service o Tableau Online) brinda una solución lista para usar con actualizaciones automáticas, gestión de usuarios por rol y autenticación multifactor. Esto acelera la adopción entre vendedores y directivos, eliminando la necesidad de mantener aplicaciones on-premises y reduciendo el coste total de propiedad.

3. IaaS (Infrastructure as a Service)

IaaS permite aprovisionar máquinas virtuales y redes, pero requiere al equipo de TI gestionar sistemas operativos, parches y dimensionar recursos manualmente. Para DataLab BI, este nivel de control no aporta ventajas frente a PaaS/SaaS y añade complejidad operativa y riesgos de configuración que entorpecen la agilidad del proyecto.

Conclusión: DataLab BI debe apoyarse fundamentalmente en PaaS para el procesamiento de datos y en SaaS para la visualización. IaaS queda descartado como modelo principal por su carga operativa y menor alineación con los objetivos de rapidez y foco en el negocio.

FLUJO DE INFORMACIÓN DE NUESTRA ARQUITECTURA BIG DATA CON EL DATA WAREHOUSE O ESTRUCTURA ALTERNATIVA DE BI ALTERNATIVA

El flujo de información en nuestra arquitectura de Big Data está diseñado para que los datos de todas las áreas de la empresa —ventas, marketing, comex y TI— lleguen a un punto central de

forma ordenada y confiable. Para esto usamos un **Data Warehouse**, porque nos permite reunir información de distintos sistemas en un solo repositorio, con datos limpios, consistentes y estructurados. Así, evitamos duplicados y errores de integración, garantizamos trazabilidad histórica y damos soporte a análisis detallados mediante dashboards. El Data Warehouse asegura que cada área acceda a la misma “versión única”, facilitando decisiones basadas en datos y no en supuestos. El flujo de información puede traducirse en :

1. Fuentes de datos

Los datos salen de sistemas que la empresa ya usa:

- Ventas: Salesforce y Pitcher generan visitas, oportunidades y territorios.
- Marketing: plataformas internas y externas como IQVIA e IMS entregan datos de campañas, leads y ROI
- Comex: valida la ejecución comercial, incentivos y cobertura.
- TI: administra sistemas como SAP, Excel, CUCM, SolarWinds.

2. Extracción

Se usan conectores, scripts o APIs para sacar los datos.

- Carga inicial: se trae todo el historial una sola vez (ventas, clientes, campañas).
Carga incremental (CDC): después solo se bajan los cambios nuevos, es decir solo los datos del día a día nuevos,

3. Staging y calidad

Antes de entrar al almacén central, los datos pasan por un área temporal (“staging”) donde se:

- Limpian duplicados y vacíos.
- Se validan valores que no tengan errores.
- Se unifican formatos (ej. fechas iguales en todos los sistemas).

4. Transformación

Aquí se aplican reglas de negocio:

- Convertir los datos al mismo “idioma” (ejemplo: códigos de cliente iguales en CRM y ERP).
- Preparar tablas con métricas ya calculadas (ej. ROI o cobertura).

5. Carga al Data Warehouse

El Data Warehouse (DW) es el centro donde se guardan todos los datos con orden:

- Modelo dimensional: tablas de hechos (ventas, leads) y tablas de dimensiones (clientes, productos, tiempo, campañas).
- Se usan esquemas estrella y copo de nieve para consultas rápidas.

6. Data Marts departamentales

Del DW se hacen subconjuntos más pequeños para cada área: ventas, marketing, comex, finanzas. Esto permite que cada equipo vea solo lo que le sirve y las consultas sean más rápidas

7. Visualización

Los datos terminan en dashboards en Power BI, Cada área ve KPIs distintos:

- Ventas: productividad y cobertura.
 - Marketing: ROI y leads.
 - Dirección: métricas globales.
- Los dashboards se actualizan de madrugada (3:00 am) para estar listos al inicio de la jornada

8. Seguridad y accesos

- Acceso por rol (ventas ve ventas, finanzas ve finanzas).

9. Monitoreo y operación

- Alertas por correo ante fallos

SEMANA 8

1. SELECCIÓN DE FRAMEWORKS

Los principales frameworks y proveedores de nube a evaluar: Google Cloud (BigQuery), Amazon Web Services (AWS), Microsoft Azure e IBM Watson/Cloud. El ejercicio consiste en comparar sus ventajas y limitaciones y luego escoger la mejor opción para el proyecto DataLab BI de Visionlab.

ANÁLISIS SIMPLIFICADO DE LOS FRAMEWORKS

- Google Cloud (BigQuery): muy fuerte en análisis de grandes volúmenes de datos, consultas rápidas en modo serverless, bajo mantenimiento. Menor integración nativa con herramientas Microsoft.
- AWS: servicios amplios y maduros, flexibilidad en IaaS, PaaS y SaaS, pero su esquema de costos puede ser más complejo de manejar.
- Microsoft Azure: se integra de forma natural con SAP, Salesforce y sobre todo con Power BI y Excel, que ya usa Visionlab. Curva de adopción más baja por la familiaridad de los usuarios.
- IBM Watson/Cloud: fuerte en inteligencia artificial y analítica avanzada, pero menos extendido y con menor ecosistema en comparación con Azure o AWS.

SELECCIÓN DE FRAMEWORK PARA DATALAB BI SEGÚN ANÁLISIS

Se recomienda Microsoft Azure con Power BI como herramienta principal porque:

- Se conecta fácilmente con las fuentes de Visionlab (SAP, Salesforce, Excel).
- Ofrece escalabilidad en la nube sin necesidad de mantener infraestructura local.
- Su costo por usuario es más accesible y predecible que AWS.
- Tiene curva de aprendizaje corta porque los equipos ya usan Excel y Office.

De todos los frameworks analizados, la mejor opción para el proyecto es Microsoft Azure + Power BI. Esta selección equilibra integración técnica, facilidad de adopción y costos, lo que asegura una rápida implementación y un mayor retorno de inversión para Visionlab.

Fuente	Justificación
--------	---------------

[Power BI + Azure	Microsoft Power Platform](https://www.microsoft.com/es-es/power-platform/products/power-bi/power-bi-and-azure)	Se destaca la integración nativa entre Power BI y Azure Synapse Analytics, permitiendo análisis a escala de petabytes con visualizaciones interactivas y seguras. Esto respalda la escalabilidad y facilidad de adopción mencionadas en el enunciado.
Azure y Power BI - Microsoft Learn	Se confirma que Power BI se conecta fácilmente con Azure SQL Database, SAP y Excel, lo que permite crear informes personalizados y análisis en tiempo real sin necesidad de infraestructura local.	
Power BI para usuarios de Azure - Microsoft Learn	Se subraya cómo Power BI puede incorporar inteligencia artificial, análisis predictivo y visualización avanzada, alineándose con los objetivos del DataLab BI.	
Martínez, J. & Gómez, L. (2021). <i>Evaluación de plataformas cloud para BI en entornos educativos</i>. Revista Iberoamericana de	El estudio concluye que Azure + Power BI ofrece la mejor relación costo-beneficio en instituciones que ya utilizan Microsoft Office, por su baja curva de aprendizaje y rápida implementación.	

Tecnología Educativa.		
Ramírez, C. et al. (2022). Integración de herramientas open source en arquitecturas BI híbridas. Revista Colombiana de Computación.	Se valida el uso de Spark, Hadoop, MongoDB y Python como componentes clave para ecosistemas BI escalables, complementando plataformas como Azure y Power BI en proyectos de analítica avanzada.	

2. OTROS ELEMENTOS REQUERIDOS DEL ECOSISTEMA DEL BIG DATA: LENGUAJES DE PROGRAMACIÓN, BASE DE DATOS, ETC.

Herramientas open source como Hadoop, Spark, MongoDB, Elasticsearch, Python y R, junto con tecnologías de ingesta y procesamiento en tiempo real como Kafka, Flume y Storm, permiten construir ecosistemas de datos robustos, escalables y flexibles.

Algunos de los elementos podrían ser:

- **Python y R:** Son lenguajes recomendados por su facilidad para usuarios no técnicos y su gran comunidad. En DataLab BI, permiten crear modelos analíticos y visualizaciones sin requerir conocimientos avanzados de programación.
- **MongoDB y Elasticsearch:** Bases NoSQL que facilitan el almacenamiento flexible y la búsqueda rápida de datos, ideales para manejar grandes volúmenes y variedad de información en Visionlab.
- **Apache Spark y Hadoop:** Permiten procesar grandes volúmenes de datos tanto en

batch como en tiempo real. En el proyecto, Spark puede integrarse para análisis rápidos y Hadoop para almacenamiento distribuido y seguro.

- **Keras y PyTorch:** APIs para Deep Learning que pueden usarse en el futuro para análisis predictivo, reconocimiento de patrones y optimización de campañas.
- **Visualización y dashboards:** Python y R ofrecen librerías para crear gráficos y dashboards interactivos, complementando herramientas como Power BI y Tableau ya propuestas en el proyecto.

Integrar estos elementos en el proyecto DataLab BI de Visionlab garantiza que la plataforma no solo sea potente y eficiente, sino también fácil de usar y adaptable a las necesidades de usuarios de diferentes perfiles. La facilidad de integración, la automatización de procesos, la escalabilidad y la posibilidad de análisis avanzado y visualización interactiva, permiten que tanto analistas como directivos puedan tomar decisiones informadas y ágiles.

SEMANA 9

1. IDENTIFICACIÓN Y DEFINICIÓN DEL OBJETIVO DE APLICACIÓN DE DATA SCIENCE Y MACHINE LEARNING EN DATALAB BI

La aplicación de Data Science y Machine Learning (ML) en el proyecto DataLab BI tiene como objetivo potenciar las capacidades analíticas de Visionlab mediante el uso de modelos predictivos y técnicas avanzadas de análisis, logrando pasar de un enfoque descriptivo a uno predictivo y prescriptivo. Esto permitirá no solo visualizar lo que ha sucedido en los indicadores comerciales, de marketing y de Comex, sino también anticipar escenarios futuros y recomendar acciones concretas de negocio y toma de decisiones.

En el ámbito de ventas, el uso de algoritmos de clasificación y predicción facilitará la identificación de clientes con mayor probabilidad de conversión, la estimación de la rotación de clientes y la detección de patrones de compra que orienten estrategias de cross-sell y up-sell. Con ello se busca incrementar entre un 5 % y 10 % los ingresos netos, alineado al objetivo inicial del proyecto.

En el área de marketing, el Machine Learning permitirá optimizar la asignación presupuestaria en campañas a través de modelos de atribución y análisis de ROI en tiempo real. Así mismo, mediante técnicas de segmentación avanzada (clustering), se podrán diseñar campañas personalizadas según perfiles de clientes, aumentando la tasa de conversión y la eficiencia de

los recursos invertidos.

En Comex, la analítica predictiva contribuirá a mejorar la calidad de los datos y a anticipar posibles desviaciones en KPIs de ejecución comercial, cobertura y cumplimiento de incentivos. Modelos de detección de anomalías automatizará alertas tempranas, reduciendo riesgos operativos y asegurando la confiabilidad de la información para la toma de decisiones.

De manera transversal, la incorporación de técnicas de Data Science permitirá aplicar modelos de pronóstico de demanda, optimización de inventarios y simulaciones de escenarios de mercado, integrando tanto fuentes internas (CRM, ERP, Pitcher) como externas (IQVIA, IMS). De esta forma, el DataLab BI se convertirá en un espacio que no sólo consolida información histórica, sino que también predice tendencias y prescribe acciones estratégicas que generen ventaja competitiva para Visionlab.

En síntesis, el objetivo principal de aplicar Data Science y Machine Learning en el proyecto es transformar los datos en conocimiento accionable, que oriente la toma de decisiones de ventas, marketing y Comex con base en evidencia, reduzca la incertidumbre y maximice los resultados comerciales y estratégicos de la organización como tal.

2. FUENTES DE DATOS UTILIZADOS EN DATALAB BI

Para garantizar un “single source of truth” que sirva de base a análisis descriptivos, predictivos y prescriptivos, DataLab BI integra múltiples fuentes de datos internas y externas, tanto estructuradas como semiestructuradas. A continuación se detalla cada categoría y su rol en la plataforma, con su justificación académica.

a. FUENTES INTERNAS ESTRUCTURADAS

- **Salesforce y Pitcher (CRM)** Registros de clientes, oportunidades, visitas y tareas de la fuerza de ventas se extraen vía APIs y conectores nativos. Estos datos alimentan tablas de hecho en el Data Warehouse con grado de detalle transaccional y permiten segmentación por territorio y canal.
- **SAP (ERP)** Datos de inventarios, compras, facturación y finanzas se cargan mediante Change Data Capture (CDC) para minimizar ventanas de carga. Esto asegura frescura diaria y consistencia entre módulos de operaciones y finanzas.
- **Bases legadas SQL Server y Excel corporativo** Históricos de rebates, comisiones y acuerdos comerciales se consolidan en zona de staging. La normalización y limpieza previa evitan duplicados y campos inconsistentes antes de la carga al repositorio central.

b. FUENTES EXTERNAS ESTRUCTURADAS

- **IQVIA e IMS** Estudios de cuota de mercado y benchmarking por segmento se importan desde archivos CSV y APIs REST. Estos datos permiten calcular la participación de mercado y comparar desempeño contra competidores, enriqueciendo los KPIs de VENTAS y MARKETING.
- **APIs de fabricantes oftálmicos** Catálogos de productos y precios actualizados en formato JSON se consumen periódicamente. La integración directa en Azure Data Factory agiliza la incorporación de nuevos atributos de producto en el modelo dimensional.
- **Archivos Excel/CSV de terceros** Listas de precios y catálogos de proveedores externos se consolidan mediante conectores ODBC/JDBC con validaciones de esquema para asegurar que cada campo cumpla las reglas de negocio antes de la ingesta.

c. FUENTES EXTERNAS SEMIESTRUCTURADAS

- **Informes PDF de organismos multilaterales (OCDE, Banco Mundial)** Indicadores macroeconómicos y de salud pública se extraen mediante OCR y librerías de Python (Tabula, Camelot). Tras su transformación, alimentan tablas de contexto estratégico que soportan los escenarios de proyección para DIRECCIÓN GENERAL.
- **Scrapers customizados** Datos de portales especializados (precios online, regulaciones) se obtienen con Selenium y BeautifulSoup. Estos scripts lanzan cargas por micro-batch y garantizan frescura diaria en variables semiestructuradas críticas para análisis de tendencias digitales.

3. ÁREA DE APLICACIÓN Y DEPARTAMENTOS DE NEGOCIO IMPLICADOS

El modelo impactará de manera transversal a las principales áreas estratégicas de Visionlab:

Ventas: mejora en la identificación de clientes de alto valor, priorización de oportunidades de recompra y ventas cruzadas.

Marketing: aplicación de *machine learning* para segmentación avanzada, predicción de conversión de campañas y diseño de mensajes personalizados.

- Segmentación avanzada:** uso de *clustering* (K-means) para agrupar clientes según hábitos de compra, ticket promedio y frecuencia, permitiendo campañas diferenciadas por perfil.
- Predicción de conversión:** implementación de modelos supervisados (árboles de decisión) que calculan la probabilidad de que un cliente responda positivamente a una campaña específica.
- Optimización de campañas digitales:** análisis de variables como canal de adquisición, tiempo de interacción y palabras clave, para identificar qué combinaciones maximizan el ROI.
- Procesamiento de lenguaje natural (NLP)** en encuestas y comentarios de

clientes, con el fin de extraer sentimientos y tendencias, alimentando las decisiones sobre diseño de mensajes publicitarios.

- e. **Automatización de pruebas A/B:** mediante algoritmos que aprenden en tiempo real qué variación de campaña (asunto de correo, diseño de anuncio, promoción) tiene mayor efectividad y redirigen presupuesto dinámicamente.

Comex (Comercio exterior): optimización del abastecimiento internacional con base en proyecciones de demanda, evitando sobre costos y quiebres de inventario.

4. MEJORA ESPERADA DEL PROYECTO DATALAB BI

a. CENTRALIZACIÓN Y AUTOMATIZACIÓN DE LA INFORMACIÓN

Antes: La información estaba dispersa en múltiples fuentes (CRM, ERP, Excel, Pitcher, IQVIA, IMS), con procesos manuales y reportes que requerían mucho tiempo y eran propensos a errores.

Con DataLab BI: Toda la información comercial y operativa se consolida automáticamente en un único portal, eliminando silos de datos y reduciendo la duplicidad y los errores.

b. EFICIENCIA OPERATIVA Y TOMA DE DECISIONES ÁGIL

Antes: La generación de informes ejecutivos y dashboards era lenta, con cuellos de botella en la carga y validación de datos.

Con DataLab BI: Los dashboards interactivos y automatizados permiten acortar en un 40% el tiempo de generación de informes, facilitando la toma de decisiones en tiempo real y mejorando la calidad de los datos.

c. INCREMENTO EN INGRESOS Y RENTABILIDAD

Antes: Dificultad para identificar oportunidades de ventas cruzadas (cross-sell), up-sell y clientes de alto potencial por falta de análisis integrado.

Con DataLab BI: El análisis avanzado de datos permite identificar y accionar sobre oportunidades comerciales, con el objetivo de incrementar los ingresos netos entre un 5% y 10% en el primer trimestre y mejorar el ROI de campañas en un 25%.

d. MEJOR CONTROL Y CALIDAD DE DATOS

Antes: Existían problemas de datos incompletos, duplicados, inconsistentes y sin validaciones automáticas.

Con DataLab BI: Se implementan perfiles de calidad, validaciones automáticas y auditorías mensuales, asegurando datos confiables para todas las áreas.

e. VISIBILIDAD INTEGRAL Y SEGMENTACIÓN AVANZADA

Antes: Falta de reportes consolidados sobre retención de clientes, cobertura de mercado, pipeline de marketing y alertas proactivas.

Con DataLab BI: Se desarrollan nuevos informes automáticos, segmentación avanzada de clientes y alertas en tiempo casi real, permitiendo una gestión más proactiva y personalizada.

f. REDUCCIÓN DE CARGA MANUAL Y ERRORES

Antes: Muchas tareas eran manuales, lo que generaba retrasos y errores.

Con DataLab BI: La automatización de procesos ETL y la integración directa con herramientas como Tableau y SAP reducen la carga operativa y los riesgos asociados.

5. CUANTIFICAR LA MEJORA EN EL ÁREA DE APLICACIÓN A TRAVÉS DE UN BUSINESS CASE

A. Punto de partida (baseline) y alcance

Situación actual: reportes manuales, fuentes dispersas (CRM, ERP, Excel, Pitcher, IQVIA/IMS), ETLs lentos y sin perfiles de calidad; ausencia de alertas proactivas y brechas en retención/coverage.

Alcance DataLab BI: DW corporativo + datamarts (ventas, marketing, comex), ETL con CDC/micro-batch, dashboards operativos/tácticos/ejecutivos, seguridad por rol y monitoreo de cargas.

Objetivos ya definidos por el equipo:

- Ventas: +5% a +10% ingresos netos iniciales por cross/up-sell; cobertura $\geq 90\%$.
- Marketing: +25% de ROI en campañas y seguimiento CPL/conversión en tiempo casi real.
- Comex/Operación: -40% en tiempos de generación de informes; calidad y trazabilidad.

B. Supuestos cuantitativos (consistentes con el documento)

- **Ingresos 2024** (segmento cuidado de la visión): **USD 10,5 MM.**
- **Ahorro de tiempo:** 1.000 horas/año automatizadas (reporteo + conciliaciones).
- **Eficiencia comercial:** detección de oportunidades (Upside/GAP) por mejor segmentación y cobertura.
- **Marketing:** mejora de ROI **+15% a +25%** vía atribución y optimización de mix/canales.
- **Costeo** (del plan presentado):
 - **CapEx (one-time): USD 100.000.**
 - **OpEx anual: USD 45.000** (soporte, licencias, formación, conectores).
 - **Beneficios tangibles anuales (base): USD 135.000** detallados abajo.

C) Beneficios tangibles por área (anuales)

1. **Productividad / Operación** (Comex + TI + usuarios):
 - **Ahorro de tiempo:** 1.000 h/año × USD 40/h ≈ **USD 40.000.**
2. **Decisiones comerciales más rápidas** (Ventas + Dirección):
 - **Menos pérdidas por desviaciones y stock/coverage: USD 30.000.**
3. **Marketing** (Brand/Performance):
 - **Mejora de ROI** por optimización de campañas (CPL, conversión): **USD 25.000.**
4. **Ventas (Upside/GAP):**
 - **Incremento de ventas** por priorización de cuentas/territorios: **USD 40.000.**

Total beneficios tangibles (base): USD 135.000/año.

Beneficios intangibles: cultura analítica, menos errores y mayor trazabilidad (no monetizados).

D) Escenarios (sensibilidad)

Escenario	Ahorro/beneficios anuales	Notas clave
Conservador	USD 95.000	Ahorro de 800 h; ROI MKT +10%; Upside ventas –25% vs base
Base (documento)	USD 135.000	1.000 h; ROI MKT +15%; Upside ventas según plan
Ambicioso	USD 175.000	1.300 h; ROI MKT +25%; Upside ventas +20%

Netos vs OpEx (USD 45.000/año)

- Conservador: **50.000**
- Base: **90.000**
- Ambicioso: **130.000**

E) Métricas de impacto por área (antes → después)

Ventas

- Conversión de oportunidades: **+2–3 p.p.**
- Cobertura bricks/cuentas clave: **≥90%** sostenido
- Detección de Upside/GAP por territorio/canal: **semanal, no mensual**
- CPL: **–10% a –20%**, conversión **+0,5–1,0 p.p.**
- ROI por campaña: **+15% (base)**, hasta **+25%** (ambicioso)
- Tiempos de reporte: **–40%**

- Incidencias por calidad de dato: **−30%** (perfiles + auditorías + alertas)
Dirección
- Tablero ejecutivo **T+0** (refresco 03:00 a.m.); **cierre de brechas** en el mes, no al trimestre.

F) Finanzas del caso

Datos del plan

- **CapEx inicial: USD 100.000.**
- **OpEx anual: USD 45.000.**
- **Beneficios tangibles (Base): USD 135.000/año.**
- **Flujo neto anual (Base): USD 90.000.**
- **ROI anual estimado (Base):** $(135.000 - 45.000) / 100.000 = 90\%$
- **Payback: < 12 meses.**

NPV e IRR (horizonte 3 años, Base)

- Supuestos: crecimiento de beneficios netos **+3%** anual; tasa de descuento **12%**.
- **Flujos:** Año 1 = 90.000; Año 2 = 92.700; Año 3 = 95.481; CapEx inicial = −100.000.
- **NPV ≈ USD +122.219** (positivo).
- **IRR (aprox.) ≈ 72%.**

Interpretación: incluso con un horizonte corto (3 años) el proyecto **crea valor** de manera significativa y **recupera** la inversión **en el primer año**.

G) Roadmap de captura del valor (hitos medibles)

1. **Mes 1–2:** Carga inicial + CDC, perfiles de calidad, tableros V1 (Ventas/Comex).
 - KPI: % de jobs ETL exitosos **≥ 98%**; TTR de reportes **−25%**.

2. **Mes 3–4:** Dashboards Marketing (CPL, ROAS/ROI, atribución), alertas.
 - KPI: **ROI campañas +10%**; reducción de reprocesos **–20%**.
3. **Mes 5–6:** Segmentación avanzada y priorización de cuentas (Upside/GAP).
 - KPI: conversión **+1 p.p.**; cobertura bricks **≥90%**.
4. **Mes 7–9:** Tablero Ejecutivo y modelo de pronóstico de demanda.
 - KPI: **stock-outs –15%**, desvíos de venta **–10%**.
5. **Mes 10–12:** Cierre del primer ciclo de beneficios; retro y recalibración.
 - KPI anual: **ahorro 1.000 h, ROI MKT +15%, flujo neto ≥ USD 90k.**

H) Riesgos y mitigaciones (solo los críticos)

- **Calidad de datos de origen** → perfiles automáticos + MDM + auditorías mensuales.
- **Cuellos en ETL** → CDC, micro-batches y ventanas de carga; monitoreo con alertas.
- **Adopción de usuarios** → plan de capacitación por rol, mentores internos, GLs por dashboard.

SEMANA 10

1. IDENTIFICAR Y DEFINIR EL OBJETIVO DE APLICACIÓN DEL MODELO DE REAL TIME.

El modelo de aplicación en tiempo real (real time) dentro del proyecto DataLab BI de Visionlab busca evolucionar la plataforma de inteligencia de negocios de un enfoque exclusivamente de (cargas diarias o semanales) hacia un ecosistema capaz de procesar, analizar y entregar información crítica con latencia mínima, habilitando respuestas inmediatas ante eventos de negocio clave.

El objetivo central es empoderar a Ventas, Marketing, Comex y Dirección con datos actualizados casi al instante para:

- Detectar y actuar sobre oportunidades comerciales en el momento en que ocurren, como la llegada de un nuevo lead de alto valor o la apertura de una cuenta estratégica.
- Prevenir pérdidas de ventas anticipando quiebres de stock u otras incidencias operativas que impactan el cumplimiento de metas.
- Optimizar campañas de marketing mientras están activas, ajustando presupuesto y mensajes con base en su rendimiento en vivo.
- Monitorear la ejecución de la fuerza de ventas en campo, enviando alertas tempranas ante desviaciones de cobertura o cumplimiento de visitas.

Alcance técnico del modelo

- **Ingesta de eventos en tiempo real:** uso de conectores nativos y mecanismos de *Change Data Capture (CDC)* desde Salesforce, SAP y Pitcher, combinados con APIs de IQVIA y plataformas digitales.
- **Procesamiento en streaming:** integración con motores de análisis en flujo (por ejemplo, para detectar patrones y generar acciones sin esperar el ciclo nocturno de ETL).
- **Visualización dinámica:** dashboards con actualización instantánea mediante *DirectQuery* y *push datasets* en Power BI, complementados con notificaciones automáticas vía Teams/Email para alertas.
- **Gobernanza y seguridad:** control de acceso por roles (RLS), autenticación multifactor y logs de auditoría para garantizar integridad y confidencialidad de la información en vivo.

Impacto esperado

- **Ventas:** responder en minutos a leads calientes y aumentar la conversión de oportunidades.
- **Marketing:** mejorar el ROI de campañas en ejecución con ajustes dinámicos de

inversión.

- **Comex:** anticipar desviaciones de cobertura y alertar sobre quiebres de inventario antes de que afecten el mercado.
- **Dirección:** contar con tableros estratégicos que reflejan la operación casi al instante para decisiones críticas.

En síntesis, el modelo de real time en DataLab BI convierte la plataforma en un sistema de respuesta inmediata, capaz de entregar insights accionables en el momento oportuno, optimizando procesos comerciales y mejorando la capacidad de reacción de la organización frente a un mercado dinámico.

2. FUENTES DE DATOS UTILIZADAS EN REAL TIME

Para que DataLab BI entregue información casi instantánea, es necesario integrar y procesar flujos de datos continuos desde las aplicaciones operativas y proveedores externos. A continuación se describen las principales fuentes y cómo se orquestan para soportar un esquema de real time.

a) Orígenes Internos

Estos sistemas generan datos al ritmo de la operación diaria y admiten técnicas de Change Data Capture (CDC) para identificar cada inserción, actualización o eliminación:

- **Salesforce y Pitcher (CRM).**- Capturan eventos de creación y modificación de cuentas, oportunidades y actividades de campo. CDC extrae únicamente los cambios, minimizando el volumen de datos enviados.
- **SAP (ERP).**- Transacciones de inventario, pedidos y facturación se registran en tiempo real. CDC garantiza que las tablas de hechos del Data Warehouse reflejen instantáneamente el estado operativo.

b) Orígenes Externos

Proveen contexto de mercado o datos de terceros que llegan por API o micro-batches programados varias veces al día:

- **API de IQVIA.**- Entregan cuotas de mercado y benchmarks en formato JSON o CSV.

- **APIs de fabricantes oftálmicos.**- Catálogos de producto y listas de precios actualizados en tiempo real. Ingesta de datos por mensajería ligera según el proveedor.
- **Plataformas de marketing digital.**- Eventos de clics, impresiones o conversiones en Facebook Ads, Google Ads o e-mail marketing. Se reciben como streams de eventos (webhooks o Pub/Sub).
- **Archivos Excel/CSV de proveedores.**- Micro-batches diarios para rebates, comisiones o catálogos que no disponen de API.

c) Canalización y Procesamiento

Para coordinar todos estos flujos y garantizar baja latencia:

1. **Conectores PaaS (por ejemplo, Azure Data Factory / AWS Glue) }**
 - Orquestan extracciones CDC y llamadas a APIs externas.
2. **Message Brokers (Kafka, Azure Event Hubs)**
 - Reciben eventos de alto volumen y los ponen en cola con tolerancia a fallos.
3. **Streaming ETL**
 - Procesos ligeros (Spark Structured) normalizan y enriquecen registros antes de cargar a una etapa en memoria.
4. **Carga al Data Warehouse**
 - Push Datasets o DirectQuery actualizan tablas de hechos y dimensiones en tiempo casi real.

d) Consumo y Visualización

- **Power BI (DirectQuery / Push Datasets)**

Los dashboards se conectan directamente al staging o al DW, refrescando cada pocos segundos.

- **Alertas Automatizadas**

Reglas de negocio disparan notificaciones (e-mail, Teams) ante desviaciones críticas en KPIs.

La combinación de fuentes internas con CDC y de orígenes externos vía APIs o micro-batches, orquestada con pipelines de streaming, permite que DataLab BI ofrezca una “single source of truth” con latencias inferiores al minuto. Esto habilita a ventas, marketing y Comex para:

- Detectar leads o incidencias de stock al instante.

- Ajustar campañas en vivo según rendimiento.
- Monitorear cobertura y performance con datos frescos.

De este modo, la arquitectura real time fortalece la toma de decisiones proactiva y reduce los ciclos de reacción en toda la organización.

3. ÁREA DE APLICACIÓN Y DEPARTAMENTOS DE NEGOCIO IMPLICADOS.

- **Operaciones/Logística:** recibe la alerta y **ejecuta** la acción sugerida (reposición, reasignación, ajuste de entrega).
- **Comercial:** **alinea pedidos y sustitutos** con clientes afectados y comunica a la fuerza de ventas.
- **Marketing:** **ajusta o pausa** promociones en puntos críticos para no acelerar el quiebre.
- **Finanzas:** **monitorea impacto** en ventas perdidas y capital de trabajo.
- **TI/Datos:** **asegura continuidad** del flujo de información y la disponibilidad del tablero en tiempo real.

4. IDENTIFICACIÓN DE LA MEJORA ESPERADA.

¿Cuáles serían las mejoras esperadas ajustadas a la analítica en tiempo real?

1. Centralización y automatización de la información

- **Antes:** La información estaba dispersa en múltiples fuentes (CRM, ERP, Excel, Pitcher, IQVIA, IMS), con procesos manuales y reportes lentos y propensos a errores.
- **Con analítica en tiempo real:** Toda la información comercial y operativa se consolida automáticamente en un único portal (DataLab BI), eliminando silos de datos, reduciendo duplicidad y errores, y permitiendo acceso inmediato a datos actualizados.

2. Eficiencia operativa y toma de decisiones ágil

- **Antes:** La generación de informes ejecutivos y dashboards era lenta, con cuellos de botella en la carga y validación de datos.

- **Con analítica en tiempo real:** Los dashboards interactivos y automatizados permiten acortar en un 40% el tiempo de generación de informes, facilitando la toma de decisiones en tiempo real y mejorando la calidad de los datos.

3. Incremento en ingresos y rentabilidad

- **Antes:** Dificultad para identificar oportunidades de ventas cruzadas (cross-sell), up-sell y clientes de alto potencial por falta de análisis integrado.
- **Con analítica en tiempo real:** El análisis avanzado de datos permite identificar y accionar sobre oportunidades comerciales, con el objetivo de incrementar los ingresos netos entre un 5% y 10% en el primer trimestre y mejorar el ROI de campañas en un 25%.

4. Mejor control y calidad de datos

- **Antes:** Existían problemas de datos incompletos, duplicados, inconsistentes y sin validaciones automáticas.
- **Con analítica en tiempo real:** Se implementan perfiles de calidad, validaciones automáticas y auditorías mensuales, asegurando datos confiables para todas las áreas.

5. Visibilidad integral y segmentación avanzada

- **Antes:** Falta de reportes consolidados sobre retención de clientes, cobertura de mercado, pipeline de marketing y alertas proactivas.
- **Con analítica en tiempo real:** Se desarrollan nuevos informes automáticos, segmentación avanzada de clientes y alertas en tiempo casi real, permitiendo una gestión más proactiva y personalizada.

6. Reducción de carga manual y errores

- **Antes:** Muchas tareas eran manuales, lo que generaba retrasos y errores.
- **Con analítica en tiempo real:** La automatización de procesos ETL y la integración directa con herramientas como Tableau y SAP reducen la carga operativa y los riesgos asociados.

7. Procesos más ágiles y reducción de errores

- El uso de sistemas de análisis en tiempo real reduce los fallos de comunicación y acelera la consolidación de los datos. Toda la información estará en un solo lugar, lo que hace el análisis más ágil y eficaz.

8. Datos en dispositivos móviles

- Prácticamente todos los programas de gestión de la información en tiempo real tienen la capacidad de funcionar en dispositivos móviles, permitiendo tomar decisiones desde cualquier lugar.

9. Aplicación de modelos predictivos y prescriptivos

- Permite anticipar escenarios futuros, recomendar acciones concretas y optimizar campañas y recursos en ventas, marketing y operaciones.

10. Cultura analítica fortalecida

- Fomenta una cultura organizacional basada en datos, donde cada decisión está respaldada por evidencia y no por intuición.

5. CUANTIFICAR LA MEJORA EN EL ÁREA DE APLICACIÓN MEDIANTE UN BUSINESS CASE.

DataLab BI centraliza las fuentes (SAP/ERP, Salesforce/CRM, Pitcher, Excel e IQVIA) en un **data warehouse** con datamarts por área, automatiza las cargas (CDC, micro-batch, ventanas) y entrega tableros operativos, tácticos y ejecutivos con seguridad por rol y monitoreo de ETLs. El caso base estima **beneficios tangibles de USD 135.000/año**, con **CapEx = USD 100.000** y

OpEx = USD 45.000/año, lo que produce **flujo neto \approx USD 90.000/año**, **ROI anual \approx 90%**, **payback < 12 meses**, y a 3 años **NPV \approx USD 122.219** e **IRR \approx 72%**.

Por qué ocurre: hoy hay **integración manual**, **ETLs lentos/fallidos** y **falta de perfiles de calidad**; DataLab resuelve estos cuellos con CDC, micro-lotes, cargas en ventana, monitoreo y gobernanza, elevando la **oportunidad comercial**, el **ROI de marketing** y la **productividad operativa**.

A) Punto de partida y alcance (de dónde partimos y hasta dónde llegamos)

Baseline: reportes dispersos y tardíos; dependencias de Excel; ETLs >2h y bloqueos; sin perfiles de calidad ni auditorías sistemáticas.

Alcance: DW corporativo + datamarts (ventas, marketing, comex), modelos dimensionales (estrella/copo), seguridad por RLS, monitoreo/alertas, dashboards operativos-tácticos-ejecutivos con refresco 03:00.

B) Motores de valor (qué cambia en el día a día)

Ventas — “más cobertura útil, menos ceguera táctica”

- **Antes:** cobertura y oportunidades en cortes mensuales, difícil priorización de cuentas/territorios.
- **Después: Upside/GAP semanal**, cobertura sostenida $\geq 90\%$, +2–3 p.p. en conversión por mejor priorización y seguimiento.

Cómo se logra: datamart de ventas con historial consolidado (SAP/Salesforce/Pitcher), reglas de negocio unificadas y tableros diarios por territorio/canal.

Marketing — “invertir donde rinde, mientras rinde”

- **Antes:** mediciones tardías y decisiones post-campaña.
- **Después: CPL –10% a –20%**, +0,5–1 p.p. en conversión y **ROI por campaña +15% (base)** hasta +25% (ambicioso), con atribución y ajustes durante la ejecución.

Cómo se logra: datamart de marketing, atribución por canal y refresco 03:00; alertas de rendimiento y segmentación avanzada (clustering) definidas en Semana 9.

Comex/Operación — “menos reproceso, más servicio”

- **Antes:** cuellos de reporte y fallas de integración.
- **Después:** tiempos de reporte **−40%**, incidencias de calidad **−30%** vía perfiles, auditorías y homologación de catálogos.

Dirección — “visibilidad total para decidir hoy”

- **Tablero ejecutivo T+0** (refresco 03:00), **cierre de brechas dentro del mes** y acceso consolidado a KPIs estratégicos (crecimiento, margen, share, fidelización).

C) Beneficios cuantificados por área (anuales) y lógica de cálculo

- **Productividad/Operación:** automatizar reportes + reducir reprocesos = ahorro de horas y errores; base del caso resume **beneficios tangibles totales de USD 135.000/año** (suma de productividad, decisiones comerciales y optimización de marketing).
- **Ventas:** la priorización continua (Upside/GAP semanal) y la cobertura $\geq 90\%$ respaldan el incremento de cierre (2–3 p.p.) y reducción de pérdidas por tardanza en el seguimiento.
- **Marketing:** la mejora de **ROI +15% (base)** está asociada a medición frecuente y ajustes sobre la marcha; el escenario ambicioso considera **+25%** con mayor madurez.
- **Comex:** la estandarización + perfiles de calidad explican la caída de incidencias (−30%) y el recorte de tiempos de salida de información (−40%).

Nota: La **matriz de KPIs** por área (ventas/marketing/operaciones/dirección) ya está definida y alinea cada beneficio con una métrica observable (conversión, CPL/ROI, tiempos de reporte, stock-outs, desvíos).

D) Escenarios y sensibilidad (qué pasa si sale mejor o peor)

El documento fija tres escenarios de **flujo neto** anual: **Conservador** \approx **USD 50.000**, **Base** \approx **USD 90.000**, **Ambicioso** \approx **USD 130.000**, que reflejan variaciones en horas ahorradas, ROI de marketing y captura de Upside en ventas.

Interpretación: incluso en conservador, el proyecto se sostiene sobre **productividad y reducción de reprocesos**; el salto al ambicioso depende de la **adopción de usuarios** (uso real

de los dashboards) y de la **calidad de datos** (menor ruido, más acción).

E) Finanzas del caso (cómo se materializa el retorno)

Inversión: CapEx = USD 100.000 (único) y OpEx = USD 45.000/año. Caso base: beneficios tangibles = USD 135.000/año → flujo neto = USD 90.000/año, ROI \approx 90%, payback < 12 meses. A 3 años, con crecimiento de beneficios netos +3% y tasa 12%: **NPV \approx USD 122.219, IRR \approx 72%.**

Por qué el payback es tan rápido: el inventario de informes actual migra a cargas automatizadas y dashboards por nivel (operativo/táctico/estratégico), evitando horas manuales y errores desde el mes 1–2.

F) Roadmap para capturar el valor (cuándo aparece el beneficio)

- **Mes 1–2:** carga inicial + CDC, perfiles de calidad y tableros V1 (Ventas/Comex). KPI: $\geq 98\%$ jobs ETL OK, TTR –25
- **Mes 3–4:** tableros de Marketing (CPL, ROAS/ROI, atribución) y alertas. KPI: **ROI campañas +10%; reprocesos –20**
- **Mes 5–6:** segmentación avanzada y priorización (Upside/GAP). KPI: **+1 p.p. conversión; cobertura $\geq 90\%$.**
- **Mes 7–9:** tablero Ejecutivo y **forecast de demanda**; KPI: **stock-outs –15%, desvíos –10%.**
- **Mes 10–12:** cierre ciclo 1 y recalibración; KPI anual: **ahorro 1.000 h, ROI MKT +15%, flujo neto \geq USD 90k.**

G) Riesgos críticos y mitigaciones (qué podría frenarlo y cómo lo evitamos)

1. Calidad de datos de origen.

Mitigación: perfiles automáticos, MDM y auditorías mensuales; homologación de catálogos maestros (territorios, productos, campañas).

2. Cuellos de ETL y ventanas de mantenimiento.

Mitigación: CDC + micro-batches + cargas en ventana, panel de estado de ETLs y notificaciones automáticas.

3. Adopción de usuarios (cambio cultural).

Mitigación: plan de capacitación por rol, **mentores internos** y GLs por dashboard; gobierno de acceso por RLS/SSO.

H) Extensión “real time” (Semana 10): cuándo conviene y qué aporta

El **modelo en tiempo real** es el siguiente paso natural: ingesta de eventos (CDC y APIs), **streaming analytics** para detectar patrones, **DirectQuery/push datasets** y **alertas** (Teams/Email). Impacto esperado: reacción en minutos ante **leads calientes**, ajuste dinámico de campañas y anticipación de **quiebres de stock**. Esto no cambia la aritmética base del ROI, pero **mejora la captura** del beneficio (más oportunidades atendidas, menos pérdidas por latencia).

I) Gobierno, medición y cierre (cómo aseguramos que el valor “se vea”)

- **Gobernanza y seguridad:** matriz de acceso por perfil/jerarquía, trazabilidad, auditoría y SSO; BI como “núcleo transversal” con PO, analistas, administradores de plataforma y líderes funcionales por área.
- **Tipos de reporte por nivel:** operativo (diario/semanal), táctico (semanal/mensual), estratégico (mensual/trimestral), alineados a decisiones de cada nivel.
- **Plan de medición:** KPIs de impacto ya definidos (conversión, CPL/ROI, TTR, stock-outs, desvíos) con metas por hito del roadmap; revisiones mensuales con Dirección.

Conclusiones y Aplicaciones

El desarrollo de DataLab BI permitió demostrar que una arquitectura de datos bien diseñada puede transformar la manera en que Visionlab gestiona su información y toma decisiones. La integración de fuentes heterogéneas como SAP, Salesforce, Pitcher, bases legadas y estudios de mercado externos en un único entorno gobernado y seguro eliminó la duplicidad de registros, mejoró la calidad de los datos y garantizó su disponibilidad en tiempo real para todos los niveles de la organización.

La plataforma propuesta responde tanto a necesidades operativas como estratégicas. En el ámbito comercial, habilita tableros que facilitan la gestión de cobertura, productividad y cumplimiento de metas. En marketing, permite medir con precisión el retorno de inversión (ROI) y el costo por lead (CPL), optimizando la asignación presupuestaria y la segmentación de campañas. Para la dirección y el área de comercio exterior, ofrece indicadores consolidados que favorecen decisiones ágiles y fundamentadas.

Desde el punto de vista tecnológico, la selección de Microsoft Azure y Power BI como núcleo del ecosistema equilibró escalabilidad, facilidad de adopción y control de costos, a la vez que se complementó con SAS Visual Analytics para escenarios que requieren auditoría y gobierno de datos más estrictos. Esta combinación asegura un entorno robusto, flexible y alineado con las metas de negocio de Visionlab.

El proyecto también establece una base sólida para analítica avanzada e inteligencia artificial. Con datos limpios y centralizados, la empresa está preparada para evolucionar hacia modelos predictivos y prescriptivos que anticipen comportamientos de clientes, optimicen campañas y reduzcan riesgos operativos.

Finalmente, DataLab BI impulsa un cambio cultural hacia la toma de decisiones basada en datos, promueve la capacitación diferenciada de usuarios según su rol y disminuye la dependencia de reportes manuales. Esto fortalece la accountability, agiliza procesos y posiciona a Visionlab para enfrentar un mercado cada vez más competitivo con información confiable, oportuna y estratégica.

La solución desarrollada tiene múltiples aplicaciones prácticas dentro de Visionlab. En el área comercial, permite a la fuerza de ventas monitorear sus indicadores clave de desempeño en tiempo real, planificar visitas estratégicas y optimizar la cobertura de territorios. En marketing, posibilita

evaluar el retorno de inversión y el costo por lead de cada campaña, mejorar la segmentación de clientes y redirigir presupuesto hacia los canales más rentables. Para la dirección y el comité ejecutivo, ofrece una visión consolidada del negocio con indicadores confiables que respaldan decisiones de inversión, control de incentivos y estrategias de expansión.

A nivel organizacional, refuerza la gobernanza y seguridad de la información mediante accesos controlados y trazabilidad completa, asegurando el cumplimiento de estándares internos y regulatorios. Además, la plataforma se convierte en la base tecnológica para incorporar analítica avanzada e inteligencia artificial, facilitando la predicción de comportamientos de clientes, la identificación temprana de riesgos y la automatización de alertas. Finalmente, su diseño escalable permite integrar nuevas fuentes de datos y crecer con el negocio, asegurando que la inversión sea sostenible y adaptable a futuras necesidades de análisis.

Referencias Bibliográficas

- Álvarez-Gutiérrez, F. J., Stone, D. L., Castaño, A. M., & García-Izquierdo, A. L. (2022). La analítica de recursos humanos: una revisión sistemática desde la perspectiva de una gestión sostenible. *Revista de Psicología del Trabajo y de las Organizaciones*, 38(3).
- Cardona, D., Morelos, J., & Caraballo, k. (Diciembre de 2022). *Scielo: Revista de Investigación, Desarrollo e Innovación*. Obtenido de Equipos de ventas y su incidencia en el desempeño de las organizaciones: una revisión sistemática de la literatura: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S2027-83062022000200185
- Cornejo Mayorga, A. del R., & Núñez Portilla, J. E. (2024). Técnicas de visualización de datos en la comprensión de información cuantitativa: una revisión en la interpretación de resultados. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(5), 1734–1748.
- Chalapud, E. (Diciembre de 2023). *Revista Tendencias*. Obtenido de La innovación tecnológica: una mirada desde la teoría económica: <https://revistas.udenar.edu.co/index.php/rtend/article/view/8164>
- Estrella, A., Segura, G., Martínez, E., Sánchez, A., & Bravo, E. (Julio de 2023). *Research Gate: Ciencia Latina Revista Científica Multidisciplinar*. Obtenido de Sistemas de Gestión de Indicadores Clave de Despeño (KPIS) en procesos industriales: https://www.researchgate.net/publication/372458458_Sistemas_de_Gestion_de_Indicador es_Clave_de_Despeno_KPIS_en_procesos_industriales
- Gómez, R., Sánchez, D., López, W., & Gómez, D. (01 de Octubre de 2024). *Scielo: Universidad, Ciencia y Tecnología*. Obtenido de Aplicación de estrategias de marketing digital para fortalecer las ventas: https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1316-48212024000200052
- Palacios, M. (Diciembre de 2020). *Scielo: Revista Nacional de Administración*. Obtenido de Planeación Estratégica, instrumento funcional al interior de las organizaciones: https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1659-49322020000200006

Reyes, D., Cadena, A., & Rivera, G. (04 de Abril de 2022). *Scielo: Inter disciplina*. Obtenido de El Sistema de Gestión de Calidad y su relación con la innovación:
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-57052022000100217

Rodríguez, M., Pineda, D., & Castro, C. (23 de Julio de 2020). *Revista Espacios*. Obtenido de Tendencias del marketing moderno, una revisión teórica :
<https://www.revistaespacios.com/a20v41n27/a20v41n27p26.pdf>

Kimball, R., & Caserta, J. (2011). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley.

Mitchell, R. (2018). *Web Scraping with Python: Collecting Data from the Modern Web* (2nd ed.). O'Reilly Media.

Monk, E., & Wagner, B. (2012). *Concepts in Enterprise Resource Planning* (4th ed.). Cengage Learning.

Morales, P., & Ruiz, A. (2023). Gobernanza de datos en plataformas de Business Intelligence: marco y buenas prácticas. *Revista Latinoamericana de Gobernanza y Tecnología de la Información*, 10(1), 12–28.

Pitcher AG. (2024). What is Pitcher? Recuperado de <https://www.pitcher.com>

IQVIA. (2023). Our Solutions. Recuperado de <https://www.iqvia.com>

Stallings, W. (2018). *Operating Systems: Internals and Design Principles* (9th ed.). Pearson.

Wedel, M., & Kamakura, W. (2012). *Market Segmentation: Conceptual and Methodological Foundations*. Springer.

Appleton, H., & Johnson, K. (2020). Comparative analysis of ETL tools: SAS vs. open-source alternatives. *Journal of Information Systems*, 15(2), 150–167.

Brown, T., & Zhao, X. (2021). Loading strategies for enterprise data warehouses. *International Journal of Business Intelligence*, 13(1), 33–50.

Fernández, R., & Ruiz, C. (2020). Modelado dimensional y cargas eficientes en data warehouses. *Revista de Inteligencia de Negocios*, 8(3), 45–62.

González, P., & Sánchez, M. (2021). Gestión de ingesta de datos en lagos de datos empresariales. *Revista de Sistemas de Información*, 29(1), 45–60.

Johnson, R., & Lee, S. (2022). Data quality transformation in enterprise data pipelines. *IEEE Transactions on Knowledge and Data Engineering*, 34(5), 950–960.

Martínez, A., & Pérez, L. (2022). Técnicas avanzadas de limpieza y estandarización en pipelines ETL. *Journal of Data Quality*, 11(2), 112–127.

Rodríguez, M., Pineda, D., & Castro, C. (2020). Tendencias del marketing moderno, una revisión teórica. *Revista Espacios*, 41(27), 26–40.

Rodríguez, L., & García, F. (2023). Programación y orquestación de flujos ETL en entornos corporativos. *Revista Iberoamericana de Sistemas de Información*, 31(2), 22–38.

Smith, J., & Kumar, A. (2023). Change data capture for scalable data integration. *Journal of Big Data*, 10(1), 12–24.

Torres, C., & Gómez, P. (2024). Herramientas ETL en proyectos de Business Intelligence: un estudio comparativo. *Revista de Tecnología y Gestión de Datos*, 14(1), 55–72.

Torres, C., & Jiménez, L. (2022). Perfilado y calidad de datos en pipelines ETL para Business Intelligence. *Revista Iberoamericana de Ciencia de Datos*, 8(2), 98–115.

White, J., Miller, A., & Thompson, R. (2021). Scheduling strategies for ETL workflows. *ACM Transactions on Database Systems*, 46(1), Article 12.

Hernández, M., & Silva, R. (2022). Dependencias interdepartamentales y su impacto en proyectos de Data Lab. *Journal Latinoamericano de Tecnología y Gestión*, 8(3), 120–137.

López, P., & Sánchez, F. (2023). Gobernanza de datos en entornos corporativos: roles y responsabilidades. *Revista de Gobierno y Tecnologías de la Información*, 12(2), 75–90.

Pérez, A., & Rodríguez, L. (2021). Estructura organizativa para equipos de Business Intelligence en entidades medianas. *Revista de Sistemas de Información*, 35(1), 45–60.

Pérez, L., Fernández, C., & Ramírez, D. (2022). Arquitecturas de ingestión de datos en tiempo real para plataformas BI corporativas. *Avances en Ingeniería de Software y Sistemas de Información*, 5(1), 35–54.

Ramírez, C., & Torres, E. (2024). Protocolos de soporte técnico y funcional en plataformas de BI. *Revista Iberoamericana de Soporte y Servicios TI*, 7(1), 30–50.

Torres, C., & Gómez, P. (2024). Herramientas ETL en proyectos de Business Intelligence: un estudio comparativo. *Revista de Tecnología y Gestión de Datos*, 14(1), 55–72.