

EVIDENCIA DE APRENDIZAJE 2– Práctica Profesionalizante I

Materia: Práctica Profesionalizante I







Año: 2025

Profesor a cargo: Dianela Accietto

Identificación del Grupo de Trabajo

Nombre del grupo: DataTeam

Equipo y Roles

-  **Líder de Proyecto:** Maia Majzum
-  **Investigadora:** Eglimar Ramirez
-  **Técnico GenAI:** Jonathan Manuel Palomenque
-  **Diseñador UX:** Jonathan Guillén
-  **Analista de Negocio:** Valentina Pich
-  **Comunicador:** Francisco Oviedo

Proyecto de solución mediante Inteligencia Artificial Generativa para análisis de reportes de visita en distribuidores PepsiCo de Argentina.

1) Resumen ejecutivo

Este proyecto aborda un problema operativo recurrente en las distribuidoras de PepsiCo: si bien el sistema de pedidos registra check-in/check-out para validar visitas de preventistas a clientes geolocalizados, esa información no se traduce en insights accionables. Los distribuidores acceden a reportes básicos que requieren trabajo manual (descargas, limpieza y cruces en Excel) para responder preguntas clave: ¿qué porcentaje de visitas fue válida por vendedor? ¿Dónde hay incumplimientos sistemáticos? ¿Cómo se relaciona la validez de la visita con la compra efectiva? ¿Existen clientes visitados sin compra a los que se les vende posteriormente por WhatsApp o teléfono?

El proyecto tiene por objetivo validar la severidad y frecuencia del dolor, el flujo actual de trabajo y los indicadores que realmente necesitan los usuarios para decidir. Para ello realizaremos 5 entrevistas con roles diversos (dueño/gerente, jefes de ventas y administrativos/back office), documentadas con citas textuales, estimación de tiempo perdido semanal y urgencia (1–10). Como resultado, entregaremos una definición precisa del problema, una lista priorizada de KPIs y listados operativos (por ejemplo, % de visitas válidas/ inválidas por vendedor; clientes con visita inválida+pedido; válida+no compra) y los requisitos de adopción de una solución que, a partir de un Excel, produzca análisis inmediatos.

2) Contexto y definición del problema

- **Situación actual:** el sistema de toma de pedidos de la preventa registra check-in/check-out válido o inválido por cliente previamente geolocalizado para determinar si el pedido se tomó en el negocio. Los distribuidores pueden ver un reporte, pero carecen de análisis y datos útiles directamente explotables (por vendedor, período, compra, etc), por lo que dependen de procesamiento manual para convertir datos en hallazgos.
- **Dolor:** no hay visibilidad clara y estandarizada de % visitas válidas/ inválidas por vendedor ni de listados prácticos (por ejemplo, *clientes con visita inválida y pedido*) que permitan priorizar acciones con rapidez. Esta brecha obliga a armar planillas, crear filtros y combinar fuentes cada semana para responder preguntas repetitivas.
- **Impacto:** el proceso actual implica consumo de tiempo en Excel, decisiones tardías por falta de reportes listos, riesgo de oportunidades perdidas (recupero de clientes sin compra, correcciones de ruta) y menor control del cumplimiento de visitas, lo que afecta la efectividad del equipo de campo y la previsibilidad comercial.
- **Oportunidad:** una aproximación “archivo-adentro/insight-afuera” permitiría que, con un Excel exportado del sistema, se calculen KPIs y se entreguen datos listos para acción (tablas y listados exportables) sin exigir manipulación adicional. Esto habilitaría revisiones semanales más ágiles y decisiones mejor informadas.

Definiciones operativas:

- **Visita válida:** condición del sistema **Es Valido = "SI"**. La app marca esta condición cuando los eventos de **check-in y/o check-out** cumplen el **radio fijo de 150 m** respecto de la **geolocalización almacenada del cliente**.



- **Visita inválida:** condición del sistema **Es Valido = "NO"**. Ocurre cuando **check-in/ check-out** quedan **fuera de 150 m**, o cuando el **cliente no posee coordenadas** previamente cargadas, o cuando hay **discrepancia** entre CI y CO (por ejemplo, **CI válido y CO inválido o viceversa**) y la app clasifica la visita como no válida. *(Se respeta exactamente lo que marca **Es Valido**.)*
- **Pedido (compra) / No compra:** **Existe Carga = "SI" / Existe Carga = "NO"**.

Parámetro: radio **150 m** (definido por la aplicación de origen).

Cálculo de métricas: se usan **las etiquetas del Excel** tal cual: válidas = **Es Valido = "SI"**, inválidas = **Es Valido = "NO"**, y compra/no compra según **Existe Carga**. **No** se crean nuevas categorías ni se excluyen filas si la app ya clasificó la visita.

Excel de prueba:

[Datos de prueba distribuidora Pampa Distribuciones S.A. Periodo del 1 al 15 de Septiembre](#)

3) Objetivos de la investigación

1. Validar urgencia y frecuencia del problema (tiempo semanal perdido, momentos de mayor fricción): cuantificar el esfuerzo que demanda transformar el reporte en información útil y en qué etapas se concentra la carga (descarga, limpieza, cruce, generación de cortes/listados).

2. Entender el flujo actual de trabajo, herramientas y barreras: relevar paso a paso cómo se usa el reporte (qué se hace, con qué herramientas) e identificar obstáculos que impiden una lectura rápida y confiable del cumplimiento.

3. Identificar KPIs y listados imprescindibles para la gestión diaria/semanal:

priorizar aquellos indicadores y tablas operativas que los usuarios realmente necesitan para decidir y que hoy demandan armado manual.

4. Medir disposición a adoptar/pagar por una solución que automatice el análisis sobre Excel: relevar urgencia (1–10), condiciones mínimas de adopción (formato de salida, exportables, filtros) y referencias de pago, a fin de evaluar la viabilidad de continuar hacia una solución en la siguiente etapa.

4) Research de usuarios — método y muestras

Metodología basada en entrevistas exploratorias. Se simulan 5 entrevistas a empleados o dueños de distribuidoras de PepsiCo Argentina (sin grabaciones).

4.1 Perfiles y cuotas

- 1 × Dueño/ Gerente
- 2 × Jefes de Ventas (distribuidores distintos)
- 1 × Administrativo / Backoffice
- 1 × Supervisor de preventistas.

4.2 Guía de consignas (15–20')

- 1)** ¿Cuál es tu rol en la distribuidora? ¿Cuántos preventistas tenes a cargo?
- 2)** Actualmente, ¿consultas el reporte de Check In Check Out que obtenes de la web de pedidos? ¿Cada cuanto lo consultas?
- 3)** ¿Cuándo fue la última vez que descargaste el Excel del reporte?



- 4) ¿Encontrás que los datos del reporte son fácilmente medibles o identificables? ¿Qué información validas diaria o semanalmente?
- 5) ¿Qué herramientas usas para trabajar en el reporte (Excel, tableros, web)?
- 6) ¿Qué información es la más necesaria para el análisis?
- 7) ¿Cuánto tiempo dedicas semanalmente a convertir o cruzar los datos para obtener información valiosa? ¿Cual es tu mayor desafío?
- 8) ¿Qué es lo más frustrante de este proceso?
- 9) ¿Cómo sería para vos la solución ideal?
- 10) En una escala del 1 al 10, ¿Qué tan urgente es resolver este problema para vos?
- 11) ¿Estarías dispuesto/a pagar por una solución? ¿Cuánto?
- 12) ¿Qué tendría que tener una solución para que la adoptes inmediatamente?
- 13) ¿Conoces alguna herramienta que pueda ayudarte en el proceso? ¿Por qué no la usas?
- 14) ¿Hay algo más que te parezca importante?

4.3 DOCUMENTACIÓN — entrevistas SIMULADAS

Según consigna, se presentan fichas simuladas: perfiles, fecha, resumen, citas textuales, dolor (1–10) y disposición a adoptar.

[ENTREVISTA 1](#)

[ENTREVISTA 2](#)

[ENTREVISTA 3](#)

[ENTREVISTA 4](#)

[ENTREVISTA 5](#)

5) PROPUESTA DE SOLUCIÓN GenAI para Análisis de reportes de visitas - Distribuidores PepsiCo

5.1 Descripción funcional:

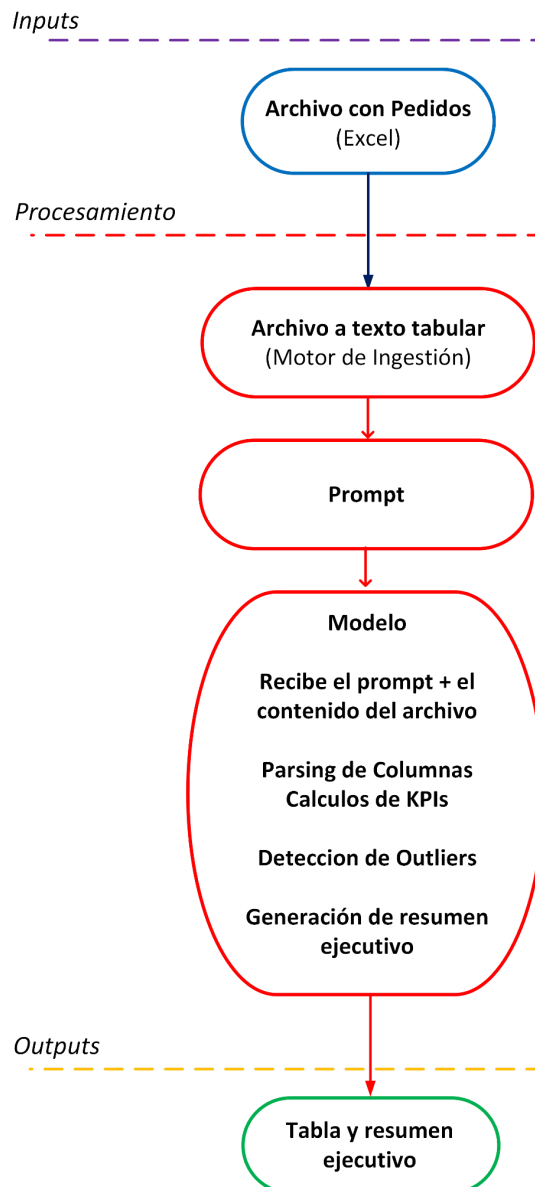
Se plantea como solución un sistema donde el usuario no tenga que hacer procesamiento manual. Este sistema permitirá cargar el Excel exportado desde el sistema de pedidos de las distribuidoras de PepsiCo y devolver métricas clave, rankings y listados accionables de manera automática, reduciendo el análisis manual a segundos

Flujo de trabajo:

- a. Usuario sube el Excel.
- b. Se envía el archivo tal cual al motor de ingestión, éste lo convierte en texto tabular (CSV/texto) para incluirlo en el prompt (sin transformaciones).
- c. El LLM recibe el prompt junto con el contenido del archivo y realiza:
 - i. Parsing de columnas y normalización básica (fechas, SI/NO).
 - ii. Cálculo de KPIs por vendedor (conteos, porcentaje de visitas válidas/ inválidas, tiempo promedio en PDV, número de casos inválidos con pedido, duplicados CI/CO, clientes sin coordenadas).

- iii. Detección de outliers (por ejemplo, vendedores con mas del 20% inválidas).
- iv. Generación de resumen ejecutivo
- d. La interfaz muestra la tabla y el resumen, con exportación disponible.

Diagrama de flujo de trabajo:



		dataset.	precisión en métricas complejas o cálculos con muchas dependencias
--	--	----------	--

Se preseleccionan los modelos de Gemini 2.5 Pro y GPT-5 ya que son los modelos más potentes del mercado y capaces de procesar grandes archivos excels.

Pruebas de rendimientos de modelos preseleccionados:

Se realizan pruebas del prompt con una excel de muestras

1. **Gemini 2.5 Pro:** a nivel general tuvo un mejor desempeño, mayor latencia a la hora de generar la respuesta. Se verificaron alucinaciones respecto a los valores de las métricas. Dificultad para trabajar con excel de muchos registros.

Gemini

2.5 Pro

consumo masivo. Tu objetivo es detectar patrones,...

Mostrar razonamiento

A continuación, se presenta el análisis de los datos de visitas de los preventistas.

Sección 1 – Tabla de resultados

Vendedor	Total visitas	% Válidas	% Inválidas	Tiempo promedio PDV	Casos inválidos con pedido	Duplicados CI/CO	Cientes sin coordenadas	Outlier (SI/NO)
13	129	85.27%	14.73%	7.02 min	8	13	2	NO
10	47	89.36%	10.64%	7.49 min	2	2	0	NO
2	31	100.00%	0.00%	5.25 min	0	4	0	NO
7	25	88.00%	12.00%	4.31 min	3	1	0	NO
3	7	85.71%	14.29%	2.01 min	1	0	0	NO
18	11	90.91%	9.09%	12.79 min	1	0	1	NO
6	1	100.00%	0.00%	8.13 min	0	0	0	NO
19	1	0.00%	100.00%	0.58 min	1	0	0	SI
8	1	100.00%	0.00%	0.45 min	0	0	0	NO
21	23	82.61%	17.39%	11.23 min	1	2	0	NO
17	1	100.00%	0.00%	13.87 min	0	0	0	NO

2. **ChatGPT-5:** brindó mejores resultados a niveles generales. Mayor tiempo de procesamiento de la respuesta, pero no generó problemas al trabajar con un excel de muchos registros.

de registro para mejorar la calidad de los datos.

He procesado el archivo y generé la tabla de KPIs por vendedor. He guardado los resultados en un Excel descargable.

- [Descargar resultados \(Excel\)](#)

A continuación te doy el **Resumen ejecutivo** (máximo 20 líneas) y algunas notas sobre la normalización y supuestos realizados.

Resumen ejecutivo:

Vendedor	total_registr	visitas_v	%_Validas	visitas_inv	%_Invalidas	promedio	validos_colicados	CI	sin_coore	Outlier
1.0	53	51	96,23	2	3,77	6,58	2	6	0	NO
10.0	43	41	95,35	2	4,65	3,91	2	4	0	NO
11.0	36	33	91,67	3	8,33	6,06	3	5	0	NO
12.0	35	33	94,29	2	5,71	4,58	2	4	0	NO
13.0	39	36	92,31	3	7,69	7,66	3	14	0	NO
14.0	46	42	91,3	4	8,7	2,71	4	13	0	NO
15.0	38	35	92,11	3	7,89	7,08	3	6	0	NO
16.0	34	33	97,06	1	2,94	5,69	1	0	0	NO
17.0	17	0	0	17	100	2,6	17	3	0	SI
18.0	43	40	93,02	3	6,98	5,68	2	4	0	NO
19.0	33	33	100	0	0	7,74	0	8	0	NO
2.0	38	37	97,37	1	2,63	7,62	1	0	0	NO
20.0	41	33	80,49	8	19,51	3,81	8	2	0	NO
21.0	35	33	94,29	2	5,71	4,52	2	2	0	NO
3.0	34	28	82,35	6	17,65	5	6	14	0	NO
4.0	40	34	85	6	15	5,02	6	0	0	NO
5.0	46	41	89,13	5	10,87	5,77	5	8	0	NO
6.0	48	45	93,75	3	6,25	4,47	3	9	0	NO
7.0	37	35	94,59	2	5,41	5,25	2	6	0	NO
8.0	44	42	95,45	2	4,55	4,49	2	8	0	NO

Se verifican resultados correctos, por lo que se decide implementar este modelo ya que generó un mejor análisis de Excel y métricas obtenidas. Tiene buena comprensión de estructuras de tablas y puede generar cálculos estadísticos y KPIs correctos

5.3 Casos de uso concretos con ejemplos:

Caso de Uso 1: Referente de PepsiCo necesita un análisis profesional de datos, para acceder a una visión general de KPIs clave (Caso de uso principal)

Definición del Prompt:

- A. Rol: Actúa como un Analista de Datos experto en gestión de preventistas y control de visitas en distribución de consumo masivo. Tu objetivo es detectar patrones, inconsistencias y oportunidades de mejora.
- B. Contexto: Recibirás como entrada el contenido completo de un archivo Excel. Este archivo contendrá registros de visitas a puntos de venta realizadas por distintos vendedores
- C. Restricciones:
 - No inventes información ni modifiques los datos.
 - Usa únicamente los valores provistos en las columnas del archivo
 - Normaliza únicamente lo necesario:
 - Fechas en Formato: YYYY-MM-DD
 - Valores binarios "SI/NO" a mayúsculas
 - En caso de valores faltantes, menciónalo en el resumen.
- D. Tarea principal: Procesar el contenido del archivo siguiendo este flujo:
 - Parsing y normalización
 - Identifica las columnas disponibles.
 - Normaliza valores básicos: fechas.
 - Cálculo de KPIs por vendedor
 - Conteo total de registros por vendedor
 - Conteo total de visitas válidas.
 - Conteo total de visitas inválidas

- Porcentaje de visitas válidas / inválidas mediante la columna “Es Valido”.
- Tiempo promedio en punto de venta (PDV) utilizando los datos de la columna “Tiempo en PDV”. Representa el resultado en minutos
- Número de casos inválidos que incluyen pedido. Usando las columnas “Es Valido” con valor “NO” y “Existe Carga”, con valor “SI”
- Cantidad de registros duplicados (CI/CO). Debes tener en cuenta registro con el mismo código y que tengan la misma fecha de check-in o check-out
- Clientes sin coordenadas.(Coordenadas CI o Coordenadas CO vacías).
- Detección de outliers
 - Señala vendedores con más del 20% de visitas inválidas.
 - Identifica cualquier otro patrón anómalo en los KPIs.(ejemplo: tiempos extremos en PDV).
- Documento Excel descargable con los resultados de los KPIs
- Resumen ejecutivo
 - Redacta un texto breve y claro con los hallazgos principales.
 - Destaca vendedores con problemas (outliers, inválidas con carga, duplicados).
 - Presenta hallazgos de manera clara y ejecutiva.

E. Formato de salida: La respuesta debe contener dos secciones claramente diferenciadas:

- Sección 1 – Tabla en excel descargable de resultados: Columnas sugeridas:

Caso de uso 2: Control de cumplimiento y Rendimiento de Preventistas

- A. Rol: Actúa como un Jefe de Ventas de una distribuidora PepsiCo en Argentina. Responsable de supervisar el rendimiento de preventistas y garantizar que las visitas a los puntos de venta (PDV) se realicen en condiciones válidas, con foco en efectividad de ventas
- B. Contexto: Actualmente, los reportes de Check-in/Check-out se descargan en Excel y requieren entre 3 y 4 horas semanales de procesamiento manual. El jefe de ventas necesita conocer, de forma rápida y clara:
- Qué vendedores cumplen mejor las visitas.
 - Quiénes tienen mayor tasa de visitas inválidas.
 - Casos de pedidos cargados en visitas inválidas (potencial incumplimiento).
 - Ranking de efectividad para tomar decisiones de capacitación y ajustes de rutas.
- C. Restricciones:
- No inventes información ni modifiques los datos.
 - Usa únicamente los valores provistos en las columnas del archivo
 - Normaliza únicamente lo necesario:
 - Fechas en Formato: YYYY-MM-DD
 - Valores binarios "SI/NO" a mayúsculas
 - En caso de valores faltantes, menciónalo en el resumen.
- D. Tarea Principal: Analizar un archivo Excel de visitas y generar:
- a. Ranking de vendedores según el porcentaje de visitas válidas.
 - b. Listado de vendedores con más de 10% de visitas inválidas.
 - c. Casos de clientes con pedido cargado en visita inválida.
 - d. Resumen ejecutivo en lenguaje natural con hallazgos clave de las métricas

E. Formato de Salida: La respuesta debe contener dos secciones claramente diferenciadas:

- Sección 1 – Tabla en excel descargable de resultados
- Sección 2 – Resumen ejecutivo: De máximo 20 líneas, debe tener un lenguaje claro, estilo ejecutivo. Destacar los puntos más críticos y hallazgos relevantes de estas métricas en relación a los vendedores.

F. Ejemplos de Salida:

Ranking vendedores según las visitas validas:

Vendedor	Visitas Totales	Visitas validas	Porcentaje de visitas válidas
15	50	47	94.0
6	40	30	75.0

Alertas vendedores con más de 10% de visitas inválidas:

Vendedor	Porcentaje de visitas inválidas
2	42.9
21	31.5

Clientes con pedido invalido:

Vendedor	Cliente	Fecha
17	Andreacola Miguel Angel (Almacen/Despensa_A)	01-09-2025
17	Andreacola Miguel Angel (Almacen/Despensa_A)	01-09-2025

Resumen:

En la última semana, el 82% de las visitas fueron válidas. El vendedor 15 lidera con 94% de cumplimiento. Se detectaron 2 casos de pedidos cargados en visitas inválidas, lo que puede indicar incumplimiento en el proceso. El vendedor 2 registra la tasa más alta de visitas inválidas (43%), lo que requiere acciones correctivas inmediatas.

Caso de uso 3: Detección de Clientes con Visita Inválida y Pedido

- A. Rol: Actúa como un Auditor de Cumplimiento Comercial en distribuidora PepsiCo. Encargado de validar que las ventas se realicen en los puntos de venta (PDV) geolocalizados correctamente, y de detectar posibles incumplimientos en la operatoria de preventistas.
- B. Contexto: Existen situaciones en las que un preventista registra una visita inválida (fuera del radio de 150 m o con discrepancia en check-in/check-out), pero igualmente se carga un pedido. Esto genera dudas sobre la transparencia del proceso y la fidelidad de los reportes. La detección temprana de estos casos es crítica para:
- Evitar prácticas indebidas (ej. carga de pedidos fuera del PDV).
 - Corregir desvíos operativos.
 - Mantener confiabilidad en los indicadores de visitas válidas.
- C. Restricciones:
- No inventes información ni modifiques los datos.
 - Usa únicamente los valores provistos en las columnas del archivo
 - Normaliza únicamente lo necesario:
 - Fechas en Formato: YYYY-MM-DD
 - Valores binarios "SI/NO" a mayúsculas

- En caso de valores faltantes, menciónalo en el resumen.

D. Tarea Principal: A partir del archivo Excel:

- Identificar todos los clientes con al menos una visita inválida y Existe Carga = SI.
- Mostrar vendedor, cliente y fecha
- Generar un breve resumen ejecutivo con la cantidad de casos detectados y vendedores más recurrentes

E. Formato de Salida: La respuesta debe contener dos secciones claramente diferenciadas:

- Sección 1 – Tabla en excel descargable de resultados
- Sección 2 – Resumen ejecutivo: De máximo 20 líneas, debe tener un lenguaje claro, estilo ejecutivo. Destacar los puntos más críticos y hallazgos relevantes de estas métricas en relación a los vendedores.

F. Ejemplos de Salida:

Casos Detectados

Vendedor	Cliente	Fecha
25	9 DE DICIEMBRE DEL 2018 (Kiosco - Open 25_A)	2025-09-10
7	Gelman Efrain (Kiosco/Maxikiosco_B)	2025-09-12

Alertas:

Vendedor	Mensaje
25	Alerta: 2 pedidos cargados en visitas inválidas, revisar operativa de ruta
7	Caso único detectado, se recomienda seguimiento puntual

Resumen:

Se detectaron 3 casos de clientes con pedidos cargados en visitas inválidas. El vendedor 25 concentra 2 de los casos y presenta una situación con duplicados CI/CO, lo que refuerza la necesidad de auditoría inmediata. El vendedor 7 tiene 1 caso aislado. La frecuencia sugiere revisar la capacitación y el control de geolocalización de estos vendedores.

5.4 Diferenciador clave vs. competencia:

Nuestra solución se diferencia de alternativas como reportes estándar del sistema, planillas de Excel personalizadas o tableros BI genéricos en varios aspectos:

- **“Archivo-adentro/Insight-afuera”:** no requiere limpieza manual ni conocimientos técnicos. Basta con subir el Excel y se generan KPIs y listados automáticamente en segundos.
- **Lenguaje natural y contexto del negocio:** el análisis se expresa en un resumen ejecutivo en español claro, con menciones específicas a vendedores, clientes y patrones relevantes.
- **Costo accesible y escalabilidad:** al ser un flujo basado en GenAI, no requiere inversión en licencias complejas de BI ni infraestructura adicional.
- **Agilidad y foco en acción:** las salidas priorizan indicadores operativos (visitas inválidas con pedido, duplicados, clientes sin coordenadas) para tomar decisiones inmediatas sin pasos intermedios.

- **Baja barrera de adopción:** diseñado para que lo pueda usar desde un dueño de distribuidora hasta un supervisor, sin necesidad de capacitaciones extensas.

5.5 Validación inicial del concepto con entrevistados:

Tras explicar a los entrevistados el concepto de solución se recogieron sus reacciones y comentarios.

Reacciones y validación del concepto:

- Alineación con necesidades:

Los cinco entrevistados confirmaron que la propuesta responde a sus principales dolores: ahorro de tiempo, automatización del cruce de datos y disponibilidad de KPIs con datos identificables para la correcta toma de decisiones.

“Sería ideal tener un reporte que agrupe, mida y compare criterios específicos sin tener que hacerlo manual.” (Agustina, Dueña)

“Si la solución me da los resultados en segundos y puedo filtrar por vendedor o fecha, la adoptaría ya mismo.” (Matías, Supervisor)

Condiciones mínimas para adopción:

- Interfaz simple, sin capacitación extensa.
- Resultados exportables y compartibles por WhatsApp/email.
- Posibilidad de ver gráficos o rankings de vendedores.
- Compatibilidad con el archivo exportado tal cual del sistema.



Percepción de valor:

La mayoría evaluó que el ahorro de 2 a 4 horas semanales justifica pagar por la solución.

Se espera que sea accesible en costo y que no requiera infraestructura adicional.

Conclusión:

Los entrevistados confirmaron la existencia del problema, y validaron la adecuación de la solución propuesta. Manifiestan entusiasmo por probar un MVP y confirman que el mayor valor estaría en la rapidez de obtener insights y en la simplicidad de uso. Esto respalda avanzar con el diseño de un prototipo funcional.