

---

# Data Warehouse para una Fundación de Acompañamiento en Duelo



Cesar Román Zúñiga.  
Bryan Alan Vargas Chávez.  
Oscar Alexander Robles Braun.

# Introducción

---



200 Voluntarios



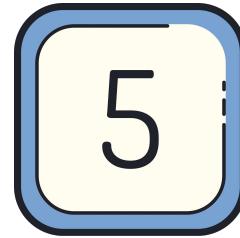
20 Ejecutivos



Donaciones y  
Alianzas



+ 1500 Dolientes



Programas de  
Acompañamiento

## Misión:

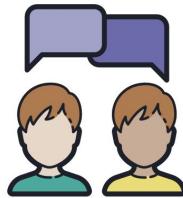
Apoyar a personas que han perdido seres queridos y promover una cultura de duelo de forma profesional

## Visión:

Crecer como fundación y ser reconocida a nivel nacional e internacional. Ser financieramente sostenible.

# Situación actual

---

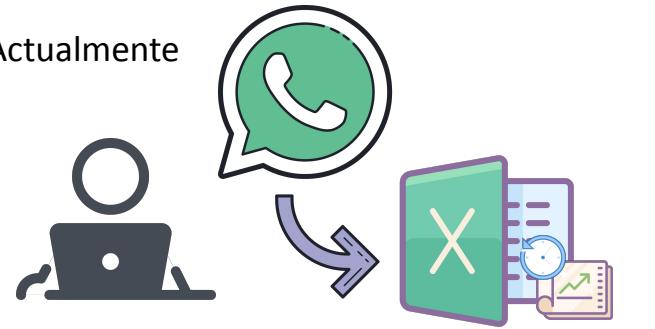


**Operación**  
Acompañamiento  
personal



**Condiciones y  
Recursos  
Tecnológicos**

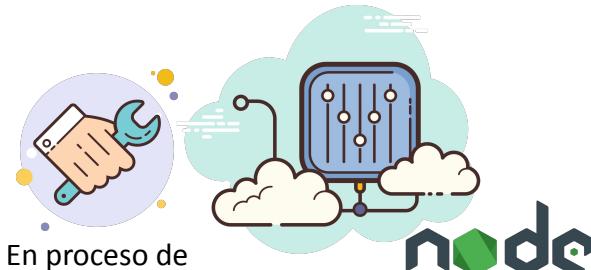
Actualmente



---

**Situación  
Actual**

A finales de 2022



En proceso de  
implementación  
(al inicio de nuestro  
proyecto)

# Objetivo

---

- El proyecto "Data Warehouse para una Fundación de Acompañamiento de Duelo" se enfocará en desarrollar un Data Warehouse que cumpla con los requisitos de la fundación en un plazo de 3 meses. Se busca categorizar al 80% de los dolientes con base a un perfil para identificar las personas que dejan el programa y permitir a la fundación tomar decisiones para disminuir el número de bajas.

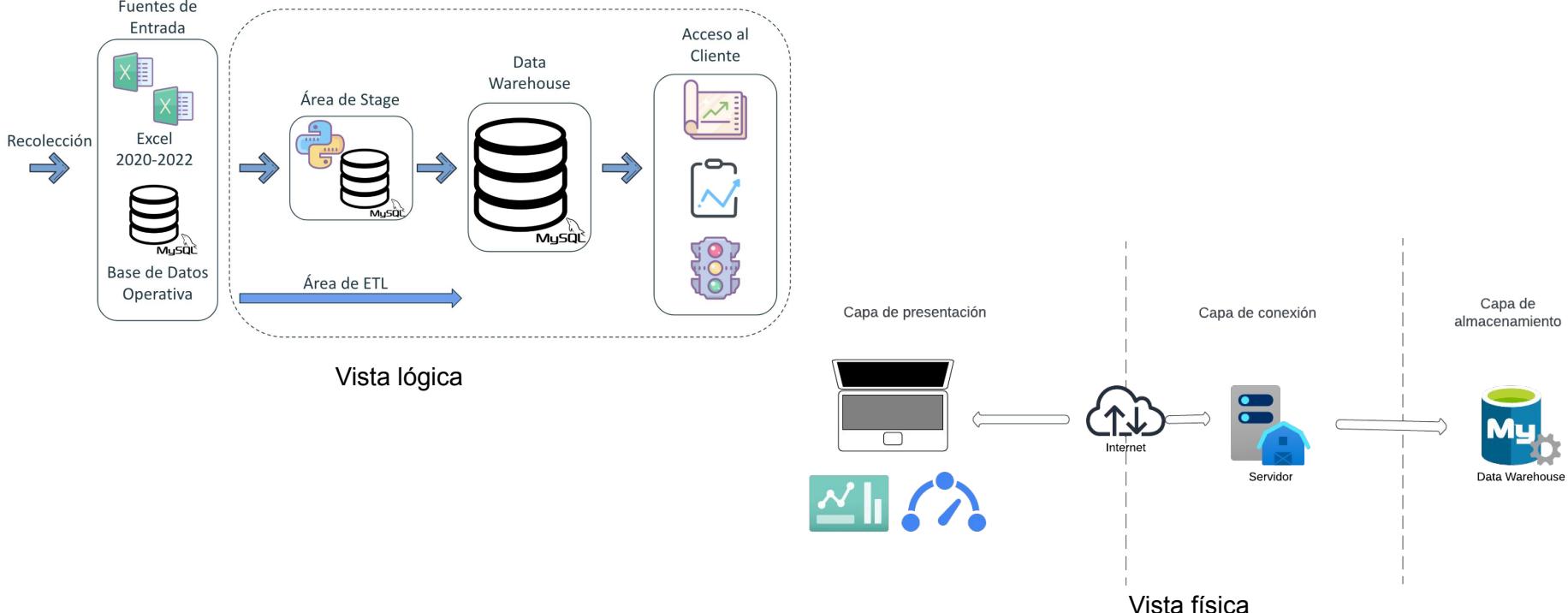


# Arquitectura: As is

---

- Al inicio del proyecto, la Fundación se encuentra **limitada** en el ámbito tecnológico. A causa del desarrollo de la Aplicación Operacional se vieron con la necesidad de **adquirir un servidor en la nube** mediante un contrato de alquiler. Gracias a este servidor y a la aplicación operativa desarrollada, la Fundación ha logrado **facilitar sus operaciones del día a día**, buscando encontrar mejores formas de apoyo para realizar sus actividades diarias.

# Arquitectura: To be



# Alcance (In)

---

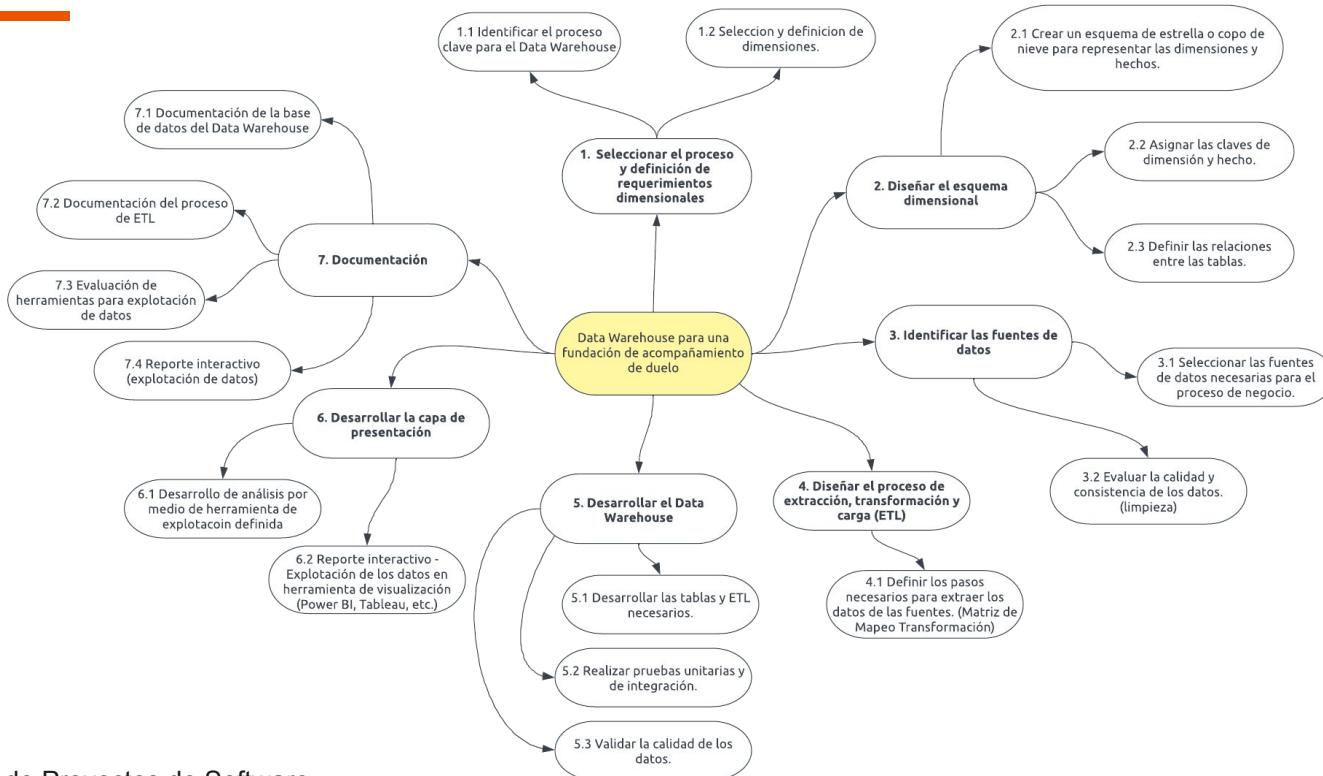
- **Diseño de un Data Warehouse:**
  - Se analizará la base de datos de la aplicación web de la fundación y la información disponible de las hojas de cálculo para modelar el Data Warehouse.
- **Limpieza de datos y ETL (extracción, transformación y carga):**
  - Se busca analizar la información qué será otorgada por parte de la fundación para su estructuración.
  - Transformar la información para su posterior carga al Data Warehouse.
- **Desarrollo de un Data Warehouse:**
  - Se utilizarán las estructuras definidas con anterioridad para desarrollar el Data Warehouse y disponer de la información.
- **Análisis de herramientas de explotación:**
  - Se considerarán las opciones para la herramienta de explotación de datos y generación de reportes en el mercado.
- **Creación de Reportes Interactivos:**
  - Se empleará alguna herramienta de análisis de datos para la creación de reportes estáticos y dinámicos.

## Alcance (Out)

---

- Este proyecto no es una continuación del desarrollo de la aplicación web.
- No se realizará la configuración, adaptación ni mantenimiento de un servidor para almacenar una base de datos.
- No se implementará ningún módulo a la aplicación web.
- No se le dará mantenimiento posterior a la entrega del Data Warehouse.

# WBS



# Costos

---

Costo por Proceso

\$8,000.00

\$6,000.00

\$4,000.00

\$2,000.00

\$0.00

Seleccionar el proceso y definición de requerimientos dimensionales.

Diseñar el esquema dimensional.

Identificar las fuentes de datos.

Diseñar el proceso de extracción, transformación y carga (ETL).

Desarrollar del Data Warehouse.

Desarrollar la capa de presentación

# **Lista de entregables**

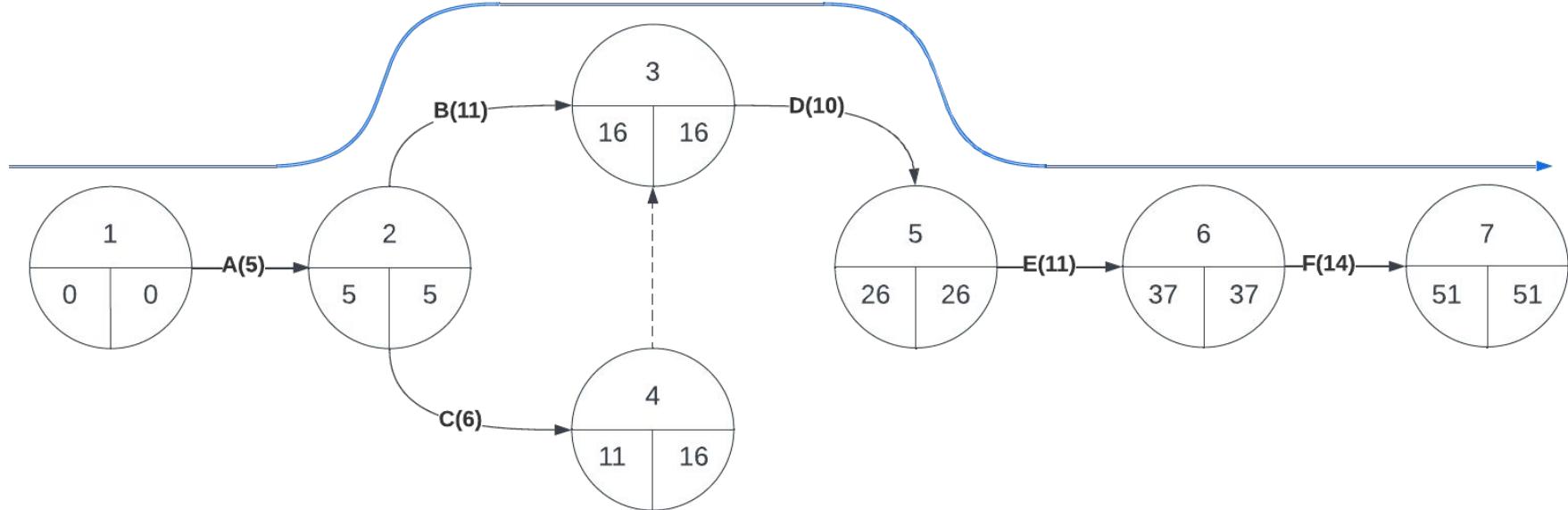
---

#	Entregable	Criterios de aceptación	Fecha
1	Documentación de la base de datos del Data Warehouse	<ul style="list-style-type: none"><li>● Utilización de la Metodología de Ralph Kimball<ul style="list-style-type: none"><li>○ Diagrama de ER</li><li>○ Explicación del diagrama</li></ul></li><li>● Metadatos</li><li>● Memoria Técnica (CREATE tablas)</li><li>● Referencias</li></ul>	14 de marzo del 2023
2	Documentación del proceso de ETL	<ul style="list-style-type: none"><li>→ Matriz de Mapeo de Transformación</li><li>→ Memoria Técnica (UPDATES/INSERTS)</li><li>→ Cumplir con al 3 de los atributos de la calidad de datos: Completitud, Validez, Exactitud<ul style="list-style-type: none"><li>- Reportes</li></ul></li></ul>	6 de abril del 2023
3	Evaluación de herramientas para explotación de datos	<ul style="list-style-type: none"><li>→ Análisis y comparación de herramientas disponibles en el mercado.</li><li>→ Definición de la herramienta más óptima para el proyecto, considerando los atributos necesarios</li></ul>	27 de abril del 2023
4	Reporte interactivo - Explotación de los datos en herramienta de visualización	<ul style="list-style-type: none"><li>→ Capacidad de visualizar el reporte según el tipo de usuario</li><li>→ Contenido de los reportes debe responder las preguntas de negocio.</li><li>→ Usabilidad de la plataforma</li></ul>	12 de mayo del 2023

# Ruta Crítica (PERT)

Actividad	Nombre del proceso	Dependencias	Pesimista	Optimista	Más Probable	Tiempo PERT
A	1. Seleccionar el proceso y definición de requerimientos dimensionales.		7	4	5	5
B	2. Diseñar el esquema dimensional.	A	12	6	9	11
C	3. Identificar las fuentes de datos.	A	8	5	6	6
D	4. Diseñar el proceso de extracción, transformación y carga (ETL).	B, C	9	7	8	10
E	5. Desarrollar el Data Warehouse	D	12	8	11	11
F	6. Desarrollar la capa de presentación	E	14	10	12	14

# Ruta Crítica (Diagrama PERT)



# Ruta Crítica (Matriz Holgura)

---

Actividad	Nodo Inicial	Nodo final	Tiempo last	Duración	Tiempo Early	Holgura
A	1	2	5	5	0	0
B	2	3	16	11	5	0
C	3	4	16	6	5	5
D	4	5	26	10	16	0
E	5	6	37	11	26	0
F	6	7	51	14	37	0

# Diagrama de Gantt

# Administración de Stakeholders

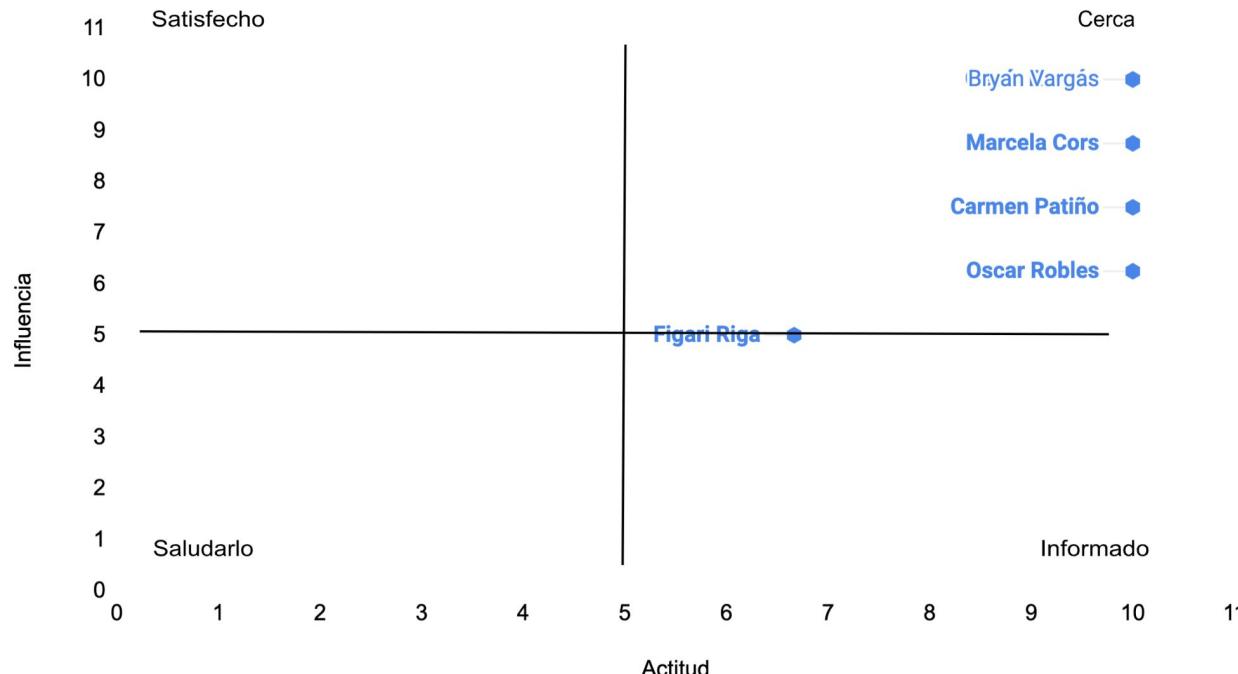
MATRIZ RACI Stakeholders			Objetivo =>	Objetivo: Proporcionar información de los datos históricos de la operación	Objetivo: Guiar al equipo de desarrolladores en cuestiones técnicas	Objetivo: Análisis y desarrollo del DWH y explotación de datos.			Objetivo: Asegurar que el proyecto del DWH se lleve a cabo de manera exitosa, considerando las necesidades del fundación			
Nombre	Dominio	Subdominio	ROL / Responsabilidades =>	R1. Conocer y explicar los metadatos de las bases de datos (Excel)	R1. Revisiones periodicas de avances del proyecto	R1. Brindar recomendaciones para un correcto desarrollo del proyecto	R1. Diseñar y desarrollar el DWH	R2. Proceso de ETL	R3. Explotación de datos	R4. Creación y entrega de la documentación del desarrollo del proyecto	R1. Dar retroalimentación del desarrollo del proyecto	R2. Mostrar el resultado final del proyecto la directora general de la fundación
Figari Riga	Operación	Operadora	Operadora	R	C	C	C	C	C	C	C	C
Carmen Patiño	Desarrollo	Asesor	Asesor del Proyecto	C	R/A	R/A	C/A	C/A	C/A	C/A	I	I
Tessy Guerrero	Desarrollo	Asesor	Asesor de base de datos	C	R	R	C	C	C	C	I	I
Bryan Vargas	Desarrollo	Desarrollador	Desarrollador de proyecto	I	I	I	R	R	R	R	I	I
Cesar Román	Desarrollo	Desarrollador	Desarrollador de proyecto	I	I	I	R	R	R	R	I	I
Marcela Cors	Gerencia	Ejecutivo	Directora General de la Fundación	A	C	C	I	I	I	I	R/A	R/A

# Administración de Stakeholders

Stakeholder	Título	Organización	Actitud	Escala	Influencia	Escala	Contribution	Escala	Actitud (nueva escala)	Influencia (nueva escala)	Recommendation (RACI)
Carmen Patiño	Asesor del Proyecto	Anahuac	Positivo	3	Fuerte	4	Soporte	2	10	7.5	Keep close
Marcela Cors	Directora General de la Fundación	Fundación Acompaña	Positivo	3	Fuerte	4	Moderado	3	10	8.75	Keep close
Figari Riga	Operadora	Fundación Acompaña	Indiferente	2	Mínima	2	Soporte	2	6.67	5	Informed
Cesár Román	Desarrollador de proyecto	Apple	Positivo	3	Fuerte	4	Recurso clave	4	10	10	Cerca
Bryan Vargas	Desarrollador de proyecto	Apple	Positivo	3	Fuerte	4	Recurso clave	4	10	10	Cerca
Oscar Robles	Desarrollador de proyecto	Apple	Positivo	3	Moderada	3	Soporte	2	10	6.25	Cerca

# Administración de Stakeholders

Actitud vs Influencia



# Administración de Riesgos

# Riesgo	Propietario del riesgo	Organización	Estatus	Riesgo escalable	Título del Riesgo	Oportunidad o amenaza	Categoría	Descripción del riesgo	Impacto potencial	Probabilidad de que ocurra el riesgo			Prioridad	Respuesta a los riesgos	Acciones de respuesta al riesgo
1	Fundación Acompaña	Fundación Acompaña	Abierto	No	Requerimientos del Proyecto	Oportunidad	Scope of Work	Los requerimientos del proyecto, son cubren las necesidades por completo	Bajo	4	Improbable	6	4	Mitigar	Analizar el esquema, para conocer si la información requerida es obtenible
2	Fundación Acompaña	Fundación Acompaña	Abierto	No	No se implemente el proyecto	Amenaza	Technician	Al final del desarrollo, el proyecto no sea implementado	Bajo	4	Probable	9	5	Aceptar	Crear una documentación con paso a paso para su implementación
3	Equipo Desarrollo	Apple	Abierto	Si	Poca experiencia con herramientas de explotación	Amenaza	Technician	El desarrollo del dashboard no tenga todas las características necesarias	Alto	12	Probable	9	9	Mitigar	Aprender a utilizar la herramienta, por medio de cursos y/o documentación
4	Equipo Desarrollo	Apple	Abierto	Si	Tiempo insuficiente para limpieza de datos	Amenaza	Technician	La limpieza de las fuentes de información sean más complejas de lo esperado	Alto	12	Improbable	6	8	Mitigar	Revisando los datos antes de realizar el PERT
5	Equipo Desarrollo	Apple	Abierto	No	Investigación de herramienta, sin opción clara	Oportunidad	Technician	No encontrar una herramienta de explotación que cumpla con las necesidades del proyecto	Medio	8	Probable	9	7	Aceptar	Seleccionar la herramienta que más se aproxime a las necesidades del proyecto
6	Equipo Desarrollo	Apple	Abierto	Si	Incompatibilidad de base de datos del desarrollo del DWH	Amenaza	Technician	El desarrollo del proyecto se realice con otra base de datos	Alto	12	Muy improbable	3	6	Mitigar	Recrear las bases de datos a la aplicación requerida
7	Equipo Desarrollo	Apple	Abierto	Si	Poca usabilidad de la visualización de datos	Oportunidad	Scope of Work	El diseño de la visualización de datos no cumple con los criterios de usabilidad	Bajo	4	Improbable	6	4	Mitigar	Elaborar el dashboard de acuerdo a UI y UX

# Conclusiones

---

## Aprendizaje de la materia:

- Plantear información relevante anticipadamente.
- Priorizar actividades importantes para el éxito del proyecto.
- Estimar tiempos de manera más precisa.

## Proyecto Practicum:

- El Diagrama de Gantt nos permitió visualizar las fechas de entrega y obtener una visión general de los resultados esperados para cada una de estas entregas.
- El diagrama de PERT nos ayudó a comprender los tiempos de las actividades y trabajar de acuerdo con un calendario para completar las tareas dentro de sus plazos correspondientes.

# Anexo

## Perfilamiento de Datos

### Inputs and Roll Ups

- [BAJAS\\_20\\_21](#)
- [DOLIENTES\\_2020](#)
- [DOLIENTES\\_2021](#)

Basic	Frequency	Masks	Quantiles	Groups
<b>Frequency Analysis</b>				
Value				
	Count	Percentage		
Enfermedad	2	0.54%		
Coronavirus	97	26.43%		
Accidente	22	5.99%		
Suicidio	16	4.36%		
Covid	10	2.72%		
enfermedad	9	2.45%		
COVID	6	1.63%		
Homicidio	5	1.36%		
Infarto	5	1.36%		
coronavirus	4	1.09%		
Paro cardiaco	4	1.09%		
accidente	3	0.82%		
Asesinato	3	0.82%		
covid	3	0.82%		
infarto	3	0.82%		
Lo asesinaron	3	0.82%		
Desaparecido	2	0.54%		
leucemia	2	0.54%		
A Julio le dio un infarto y Demetrio Neumonia	1	0.27%		
a mi hijo lo mataron, Y mi hermano de Coronavirus	1	0.27%		
Aborto	1	0.27%		
accidente automovilistico	1	0.27%		
Accidente automovilístico	1	0.27%		
accidente, enfermedad	1	0.27%		
Asesinato	1	0.27%		
Ambos coronavirus	1	0.27%		
Aún no nacían	1	0.27%		
Broncoaspiración por sobredosis	1	0.27%		
Calda, tal vez Infarto	1	0.27%		
cáncer	1	0.27%		
Cancer de páncreas	1	0.27%		
Cancer, depresión, y covid	1	0.27%		
Caír cadena por caída	1	0.27%		
Cirugía de corazón	1	0.27%		
Complicaciones de la diabetes y que no lo recibían en ningún hospital por el covid.	1	0.27%		
Corazón	1	0.27%		
Corona virus	1	0.27%		
Corona virus	1	0.27%		
Covid 19	1	0.27%		
COVID 19	1	0.27%		
COVID-19	1	0.27%		
De mi abuela y tía coronavirus y de mi esposo asesinato	1	0.27%		
Dejo de respirar	1	0.27%		

### Inputs and Roll Ups

- [Canalizados20\\_21](#)
- [Dolientes20\\_21](#)

### Dolientes20\_21

#### Columns

Expression	Type	Nulls	Not nulls	Distinct	Unique	Min	Max	Median
<a href="#">id_doliente</a>	STRING	0	1,139	1,139	1,139	1	999	485
<a href="#">marca_temporal</a>	STRING	0	1,139	373	67	1990-01-01	2021-12-30	2021-03-08
<a href="#">edad</a>	STRING	0	1,139	67	1	-1	85	43
<a href="#">ciudad_pais</a>	STRING	0	1,139	51	9	Aguascalientes	Zacatecas	Estado de MÁSxico
<a href="#">preferencia_de_horario</a>	STRING	0	1,139	4	0	AM	PM	PM
<a href="#">medio_de_enterarse</a>	STRING	0	1,139	6	2	LÁnea Origen	Soy peregrino de Magdala	No Recopilado
<a href="#">quieres_recibir_info</a>	STRING	0	1,139	1	0	FALSE	FALSE	FALSE
<a href="#">ser_querido_fecha_muerte</a>	STRING	0	1,139	584	323	1931-01-29	2023-10-30	2020-11-15
<a href="#">ser_querido_tipo_relacion</a>	STRING	0	1,139	5	0	Amistad	RelaciÁn Amorosa	Familia Directa
<a href="#">ser_querido_motivo_muerte</a>	STRING	0	1,139	8	0	Accidentes	Problemas de salud mental	Enfermedad
<a href="#">ser_querido_edad_muerte</a>	STRING	0	1,139	99	4	-1	99	5
<a href="#">genero</a>	STRING	0	1,139	3	0	Hombre	No Recopilado	Mujer
<a href="#">aporta</a>	STRING	0	1,139	2	0	FALSE	TRUE	FALSE

### Inputs and Roll Ups

#### Frequency Analysis

Value	Count	Percentage
Enfermedad	442	38.81%
COVID-19	402	35.29%
Accidentes	86	7.55%
Otras causas	61	5.36%
Problemas de salud mental	56	4.92%
Homicidio	51	4.48%
Complicaciones del parto	34	2.99%
No Recopilado	7	0.61%

# Anexo

## Perfilamiento de Datos

### Inputs and Roll Ups

- [BAJAS\\_20\\_21](#)
- [DOLIENTES\\_2020](#)
- [DOLIENTES\\_2021](#)

Basic Frequency Domains Masks Quantiles Groups

#### Mask Analysis

Value	Count	Percentage
LLLLLL DD LL LLLL	34	9.26%
LLLLLL DD LL LLLL	19	5.18%
LLLLLL DD LL LLLL	17	4.63%
LLLLLL DD LL LLLL	16	4.36%
LLLLLL D LL LLLL	14	3.81%
LLLLLL DO LL LLLL	13	3.54%
LLLLLL DD LL LLLL	13	3.54%
LLLLLL DD LL LLLL	12	3.27%
LLLL DD LL LLLL	11	3.00%
LLLL D LL LLLL	10	2.72%
LLLL DD LL LLLL	10	2.72%
LLLL D LL LLLL	10	2.72%
LLLL DD LL LLLL	10	2.72%
LLLL D LL LLLL	9	2.45%
LLLL DD LL LLLL	9	2.45%
LLLL DD LL LLLL	8	2.18%
LLLL DD LL LLLL	8	2.18%
LLLL D LL LLLL	8	2.18%
LLLL DD LL LLLL	8	2.18%
LLLL DD LL LLLL	7	1.91%
LLLL D LL LLLL	7	1.91%
LLLL D LL LLLL	6	1.63%
LLLL DD LL LLLL	6	1.63%
LLLL D LL LLLL	6	1.63%
LLLL D LL LLLL	5	1.36%
LLLL DD LL LLLL	5	1.36%
LLLLLL DD LL LLLL	5	1.36%
LLLLLL DD LL LLLL	4	1.09%
LLLLLL D LL LLLL	4	1.09%
LLLLLL DD LL LLLL	4	1.09%
LLLLLL D LL LLLL	3	0.82%
LLLLLL DD LL LLLL	3	0.82%
LLLLLL D LL LLLL	3	0.82%
LLLLLL D LL LLLL	3	0.82%
LLLLLL D LL LLLL	3	0.82%
LLLLLL DD LL LLLL	2	0.54%
LLLL DD LL LLLL	2	0.54%
LLLL D LL LLLL	2	0.54%
LLLL LL LL LLLL	2	0.54%
LLLL DD LL LLLL	2	0.54%
LLLL DD LL LLLL	2	0.54%
LLLLLL D LL LLLL	2	0.54%

### Inputs and Roll Ups

- [Canalizados20\\_21](#)
- [Dolientes20\\_21](#)

### Columns

Expression	Type	Nulls	Not nulls	Distinct	Unique	Min	Max	Median
<a href="#">id_doliente</a>	STRING	0	1,139	1,139	1,139	1	999	485
<a href="#">marca_temporal</a>	STRING	0	1,139	373	67	1990-01-01	2021-12-30	2021-03-08
<a href="#">edad</a>	STRING	0	1,139	67	1	-1	85	43
<a href="#">ciudad_pais</a>	STRING	0	1,139	51	9	Aguascalientes	Zacatecas	Estado de MÁxico
<a href="#">preferencia_de_horario</a>	STRING	0	1,139	4	0	AM	PM	PM
<a href="#">medio_de_enterarse</a>	STRING	0	1,139	6	2	LÄnea Origen	Soy peregrino de Magdala	No Recopilado
<a href="#">quieres_recibir_info</a>	STRING	0	1,139	1	0	FALSE	FALSE	FALSE
<a href="#">ser_querido_fecha_muerte</a>	STRING	0	1,139	584	323	1931-01-29	2023-10-30	2020-11-15
<a href="#">ser_querido_tipo_relacion</a>	STRING	0	1,139	5	0	Amistad	RelaciÃn Amorosa	Familia Directa
<a href="#">ser_querido_motivo_muerte</a>	STRING	0	1,139	8	0	Accidentes	Problemas de salud mental	Enfermedad
<a href="#">ser_querido_edad_muerte</a>	STRING	0	1,139	99	4	-1	99	5
<a href="#">genero</a>	STRING	0	1,139	3	0	Hombre	No Recopilado	Mujer

Basic Frequency Domains Masks Quantiles Groups

#### Mask Analysis

Value	Count	Percentage
DDDD-DD-DD	1139	100.00%

# Anexo

## Perfilamiento de Datos

**Inputs and Roll Ups**

- [BAJAS\\_20\\_21](#)
- [DOLIENTES\\_2020](#)
- [DOLIENTES\\_2021](#)

**DOLIENTES\_2020**

**Columns**

Expression		
<b>BAJAS</b>		
0		
Por favor danos tu nombre solo primer nombre.		
Qué edad tienes solo números.		
sexo		
<a href="#">Basic</a>	<a href="#">Frequency</a>	<a href="#">Domains</a> <a href="#">Masks</a> <a href="#">Quantiles</a> <a href="#">Groups</a>

**Frequency Analysis**

Value	Count	Percentage
267	64.34%	
mujer	112	26.99%
hombre	19	4.58%
MUER	15	3.61%
Hombre	1	0.24%
Mujer	1	0.24%

---

**Inputs and Roll Ups**

- [BAJAS\\_20\\_21](#)
- [DOLIENTES\\_2020](#)
- [DOLIENTES\\_2021](#)

**DOLIENTES\_2021**

**Columns**

Expression		
Field_0		
0		
Por favor danos tu nombre solo primer nombre.		
sexo		
<a href="#">Basic</a>	<a href="#">Frequency</a>	<a href="#">Domains</a> <a href="#">Masks</a> <a href="#">Quantiles</a> <a href="#">Groups</a>

**Frequency Analysis**

Value	Count	Percentage
546	75.31%	
mujer	161	22.21%
hombre	18	2.48%

**Inputs and Roll Ups**

- [Canalizados20\\_21](#)
- [Dolientes20\\_21](#)

**Dolientes20\_21**

**Columns**

Expression								
	Type	Nulls	Not nulls	Distinct	Unique	Min	Max	Median
<b>id_doliente</b>	STRING	0	1,139	1,139	1,139	1	999	485
<b>marca_temporal</b>	STRING	0	1,139	373	67	1990-01-01	2021-12-30	2021-03-08
<b>edad</b>	STRING	0	1,139	67	1	-1	85	43
<b>ciudad_pais</b>	STRING	0	1,139	51	9	Aguascalientes	Zacatecas	Estado de MÃ‰xico
<b>preferencia_de_horario</b>	STRING	0	1,139	4	0	AM	PM	PM
<b>medio_de_enterarse</b>	STRING	0	1,139	6	2	LÃnea Origen	Soy peregrino de Magdala	No Recopilado
<b>quieres_recibir_info</b>	STRING	0	1,139	1	0	FALSE	FALSE	FALSE
<b>ser_querido_fecha_muerte</b>	STRING	0	1,139	584	323	1931-01-29	2023-10-30	2020-11-15
<b>ser_querido_tipo_relacion</b>	STRING	0	1,139	5	0	Amistad	RelaciÃn Amorosa	Familia Directa
<b>ser_querido_motivo_muerte</b>	STRING	0	1,139	8	0	Accidentes	Problemas de salud mental	Enfermedad
<b>ser_querido_edad_muerte</b>	STRING	0	1,139	99	4	-1	99	5
<b>genero</b>	STRING	0	1,139	3	0	Hombre	No Recopilado	Mujer
<b>aporta</b>	STRING	0	1,139	2	0	FALSE	TRUE	FALSE

**Basic** **Frequency** **Domains** **Masks** **Quantiles** **Groups**

**Frequency Analysis**

Value	Count	Percentage
Mujer	958	84.11%
Hombre	140	12.29%
No Recopilado	41	3.60%