

ANÁLISIS PREDICTIVO DE INFECCIONES POST-TRASPLANTE HEPÁTICO MEDIANTE TÉCNICAS DE MACHINE LEARNING Y VISUALIZACIÓN INTERACTIVA

Proyecto Ciencia de Datos
2025-20

Daniel Esteban Aguilera Figueroa - María Alejandra Pérez Petro
Diego Felipe Carvajal Lombo - Jesús Manuel Ospino Bernal



Índice

- 01** Definición de problemática
- 02** Procesos actuales y desafíos
- 03** Objetivos
- 04** Arquitectura
- 05** Power Apps
- 06** Power BI
- 07** Clustering
- 08** Clasificación
- 09** Conclusiones



Definición de problemática

La Fundación Santa Fe de Bogotá es una institución de salud.

Atención a pacientes con enfermedades hepáticas - **trasplantes de hígado.**

Actualmente, la organización busca comprender la incidencia y los factores de riesgo asociados a las infecciones postrasplante hepático.

Complicación que afecta hasta el 71% de los pacientes en los seis meses posteriores al procedimiento (Tezcan et al., 2023).

El objetivo: identificar patrones, diseñar estrategias preventivas y proponer intervenciones clínicas diferenciadas según los perfiles de riesgo identificados.



Procesos actuales y desafíos



Fragmentación de datos



Falta de estandarización



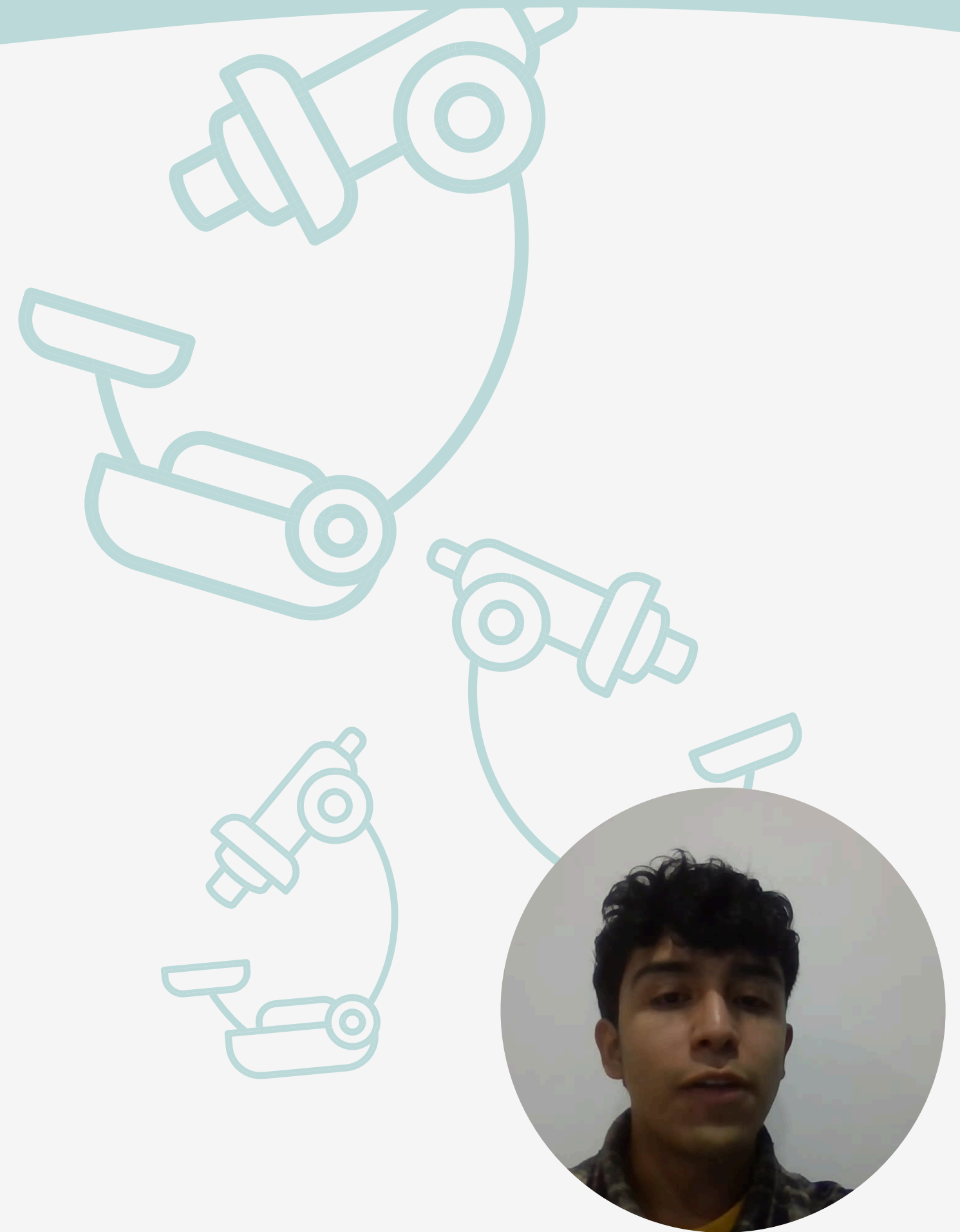
Ausencia de trazabilidad



Baja calidad de datos

Objetivo del proyecto

Aprovechar las técnicas de ciencia de datos para analizar, estructurar y optimizar la información clínica de los pacientes sometidos a trasplante hepático en la Fundación Santa Fe de Bogotá, con el fin de identificar factores de riesgo asociados a infecciones post-trasplante y fortalecer la toma de decisiones clínicas y administrativas a partir de evidencia basada en datos.



Productos

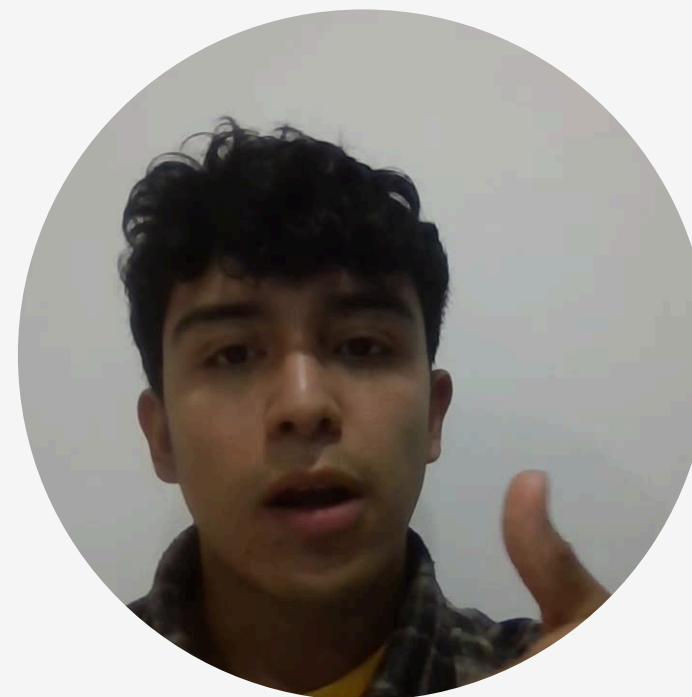


Proceso de recolección de datos utilizando herramientas de PowerApp

Aplicación digital para la captura estandarizada de información de pacientes trasplantados.

Personal médico y administrativo encargado de registro

Formularios estandarizados, integración de bases, guía técnica



Tablero Power BI

Herramienta interactiva de visualización que integra indicadores clave (KPIs) permitiendo monitorear la evolución clínica y apoyar decisiones estratégicas del hospital.

Jefe de departamento, investigadores y médicos tratantes

Conexión tiempo real, filtros dinámicos, generación reportes

Productos



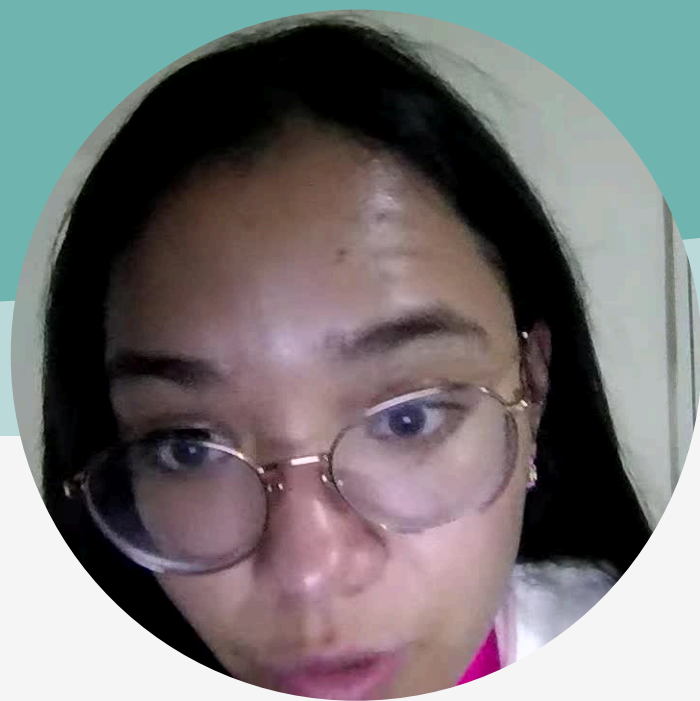
Modelo de Machine Learning

Modelo de clustering (perfilamiento) + Modelo de clasificación binaria en infección post-transplante

Equipo médico e investigadores del área de hepatología

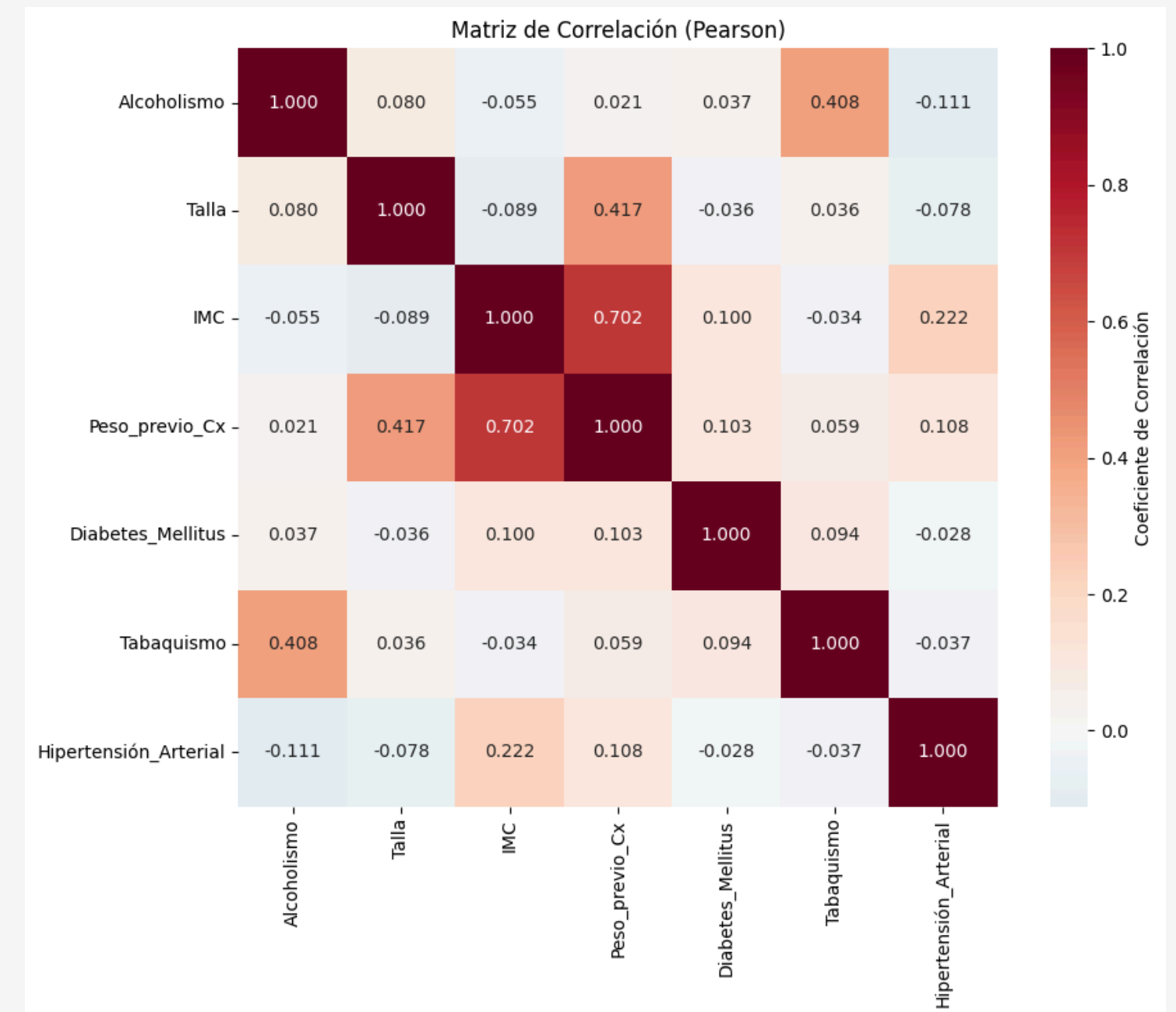
Entrenamiento modelo
Clasificación nuevos pacientes
Documentación





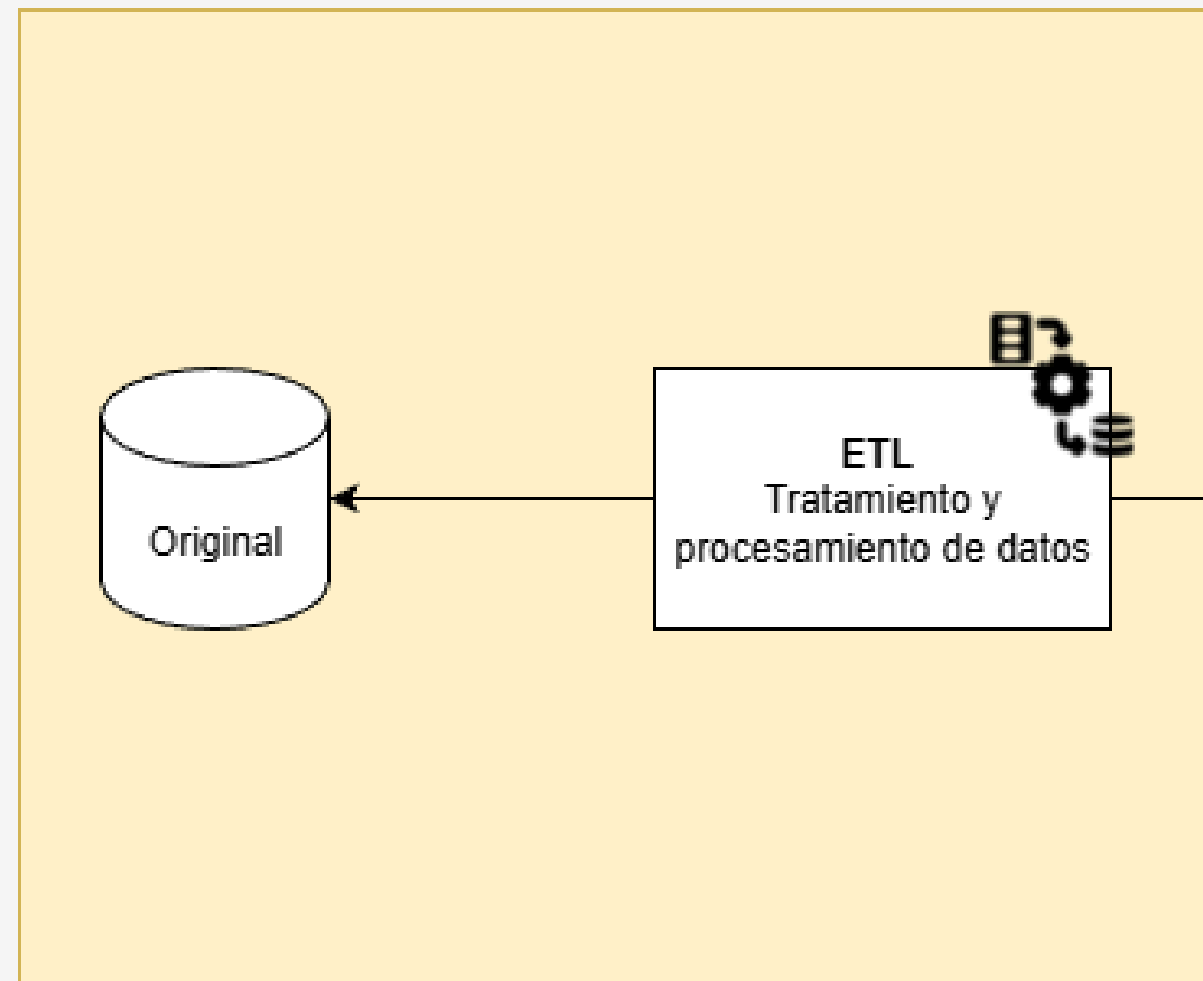
Análisis y Limpieza de Datos

- El 51% de los pacientes son hombres mientras que el 49% son mujeres.
- El 35% reporta alcoholismo, 30,8% tabaquismo, el 29,4% diabetes, 19,3% hipertensión arterial.
- La distribución de la edad se caracteriza por tener una media en 55 años, pero una oblicuidad con tendencia a valores mayores, por lo que los datos se agrupan en edades más altas, lo que es consistente con lo que se esperaba de estos.
- La media de la estatura fue de 160 cm y del peso previo al trasplante de 68 kg.
- El 17,2% de los pacientes presentó infección al momento del trasplante, mientras que el 82,8% no.
- Hay registro de 27 enfermedades (etiología) distintas que conllevaron a la enfermedad hepática siendo la más prevalente NASH con un 25,7% de los pacientes.

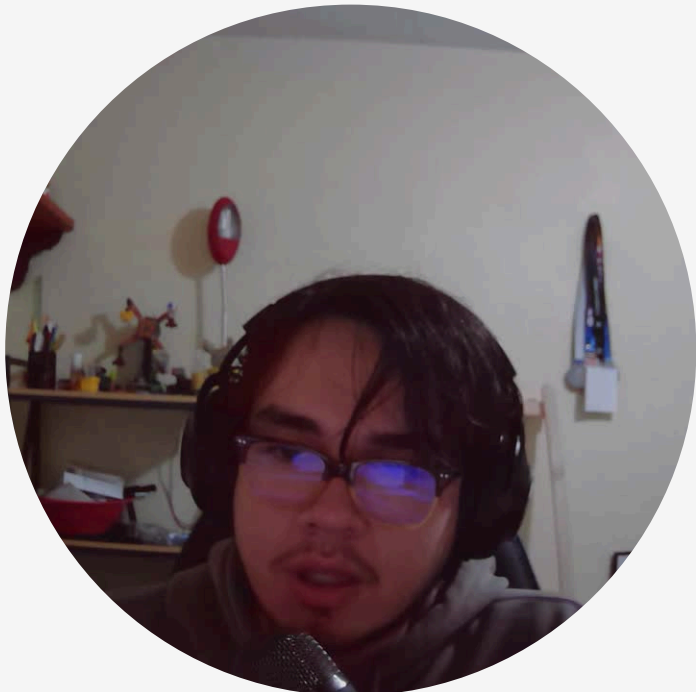
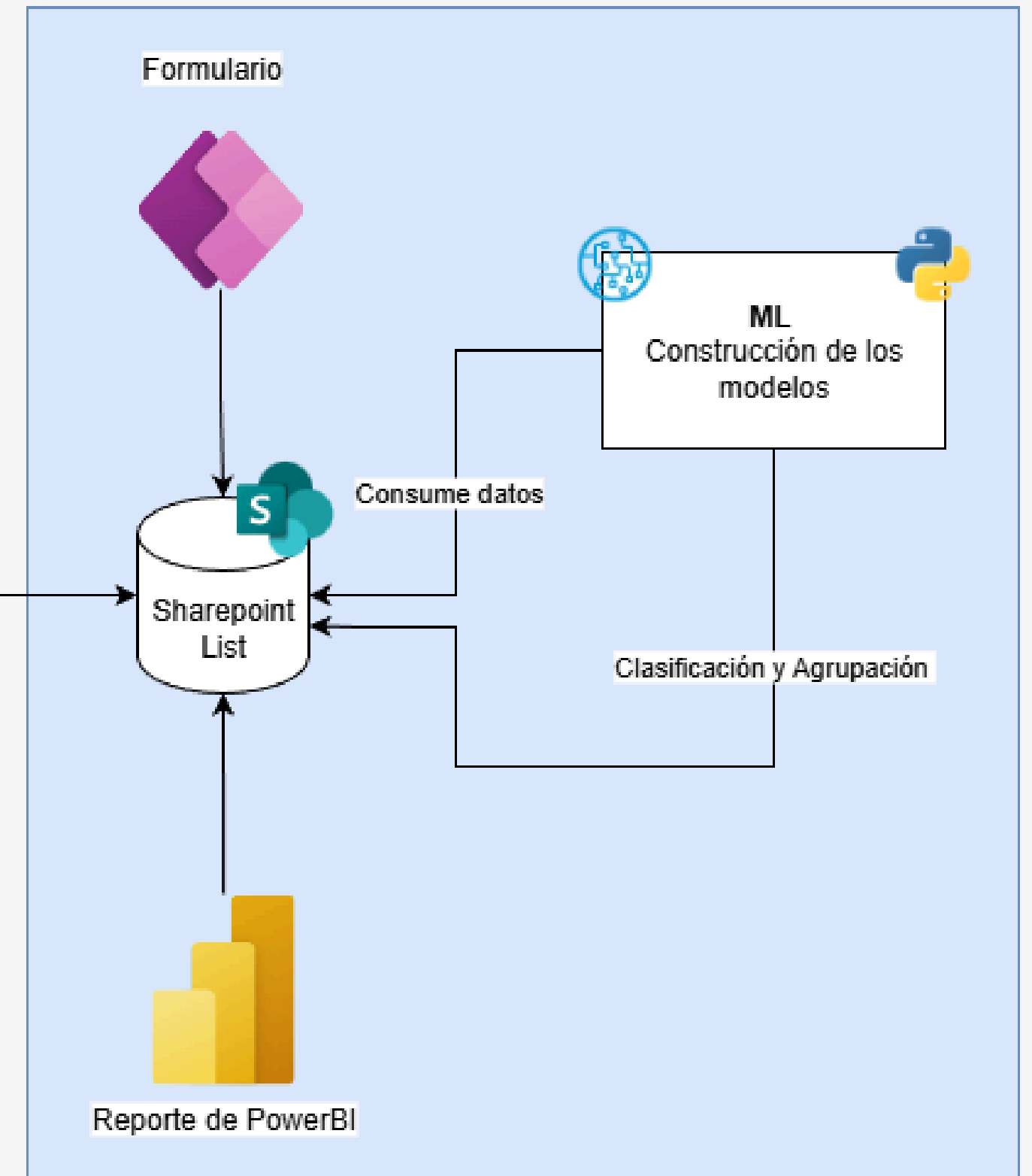


Arquitectura

Tratamiento inicial



Ambiente actual



Power App

Power Apps | Soluciones

Maria Alejandra Perez Pr

Trasplantes Fundación S

apps.powerapps.com/play/e/default-fabd047c-ff48-492a-8bbb-8f98b9fb9cca/a/5a83b0b6-6181-41dc-a420-82

Universidad de los Andes

Power Apps | Trasplantes Fundación Santa Fé

Compartir

Diego Felipe Carvajal Lombo

Actualizar y registrar pacientes con trasplantes

Diego Felipe Carvajal Lombo

Buscar

+ Nuevo

Editar

Borrar

P001
62
M

P002
62
F

P003
56
F

P004
58
M

P005
62
F

P006
66
M

P007
61

1 - 25 de 500+

← →

Página 1

#Paciente_Tx

1

Edad

62

Child_Pugh_Score_Cat

DESCONOCIDO

Fecha_Tx

vi., 2 de ene. de 2009

Alcoholismo

SI

Antibiótico_1_Tx#1

NO

Antibiótico_3_Tx#1

Número de identificación

P001

Sexo

M

Meld_Score

18

Talla

170

Antibiotico_previo_al_Tx

NO

Antibiótico_2_Tx#1

GENTAMICINA

Antibiótico_Profiláctico_VS_Terapéutico

Cancelar

Enviar

24°C
Mayorm. nublado

Buscar

10:06 p. m.
30/11/2025

Power BI

FileHomeInsertModelingViewOptimizeHelp

Get dataExcelOneLakeSQL ServerEnter dataData

Transform dataRefresh dataQueries

New visualText boxMore visualsInsert

New visual calculationNew measureQuick measureCalculations

SensitivitySensitivity

PublishShare

Prep data for Copilot AICopilot

Información individual por código de anonimización

Última vez actualizado | 23/11/2025

Código de anonimización

Sexo

Edad

Peso previo (KG)

Talla

IMC

Fecha del transplante

p001

M

62

75

170

25.95

1/2/2009 9:00:00 PM

Factores de riesgo

Diabetes ✖ · HTA ✖ · Tabaquismo ✔ · Alcoholismo ✔

Información hepática

| Etiología_#1 | Etiología_#2 | Child_Pugh_Score | Meld_Score |
|--------------|--------------|------------------|------------|
| HCV | NINGUNA | -1.00 | 18.00 |

Información previa al transplante

| UCI_al_Momento_Tx | Infección_Al_Momento_Tx | Antibiotico_previo_al_Tx | UCI_6_meses_Pre_Tx |
|-------------------|-------------------------|--------------------------|--------------------|
| NO | NO | NO | NO |

Post-operatorio

| Tiempo_Cx | Reintervención | Días en UCI | Días en el hospital | Vivo_Hoy |
|-----------|----------------|-------------|---------------------|----------|
| 8.67 | SI | 3.00 | 7.00 | NO |

Infecciones

| Germen/Microorganismo aislado | Localización de la infección | Antibiótico_Profiláctico_VS_Terapéutico |
|-------------------------------|------------------------------|---|
| DESCONOCIDO | DESCONOCIDO | Profiláctico |

Vista individualIndicadores de DesempeñoAnálisis de Resultados de Pacientes Traspla...Análisis de MuestrasPágina 1

FiltersVisualizationsData

Build visual

Filters on this pageAdd data fields here

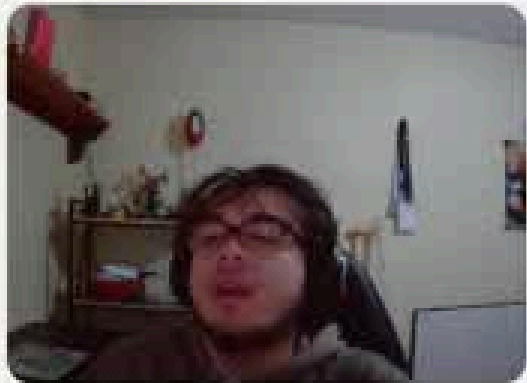
Filters on all pagesAdd data fields here

ValuesAdd data fields here

Drill throughCross-reportKeep all filtersAdd drill-through fields here

dim measureBase microbiologíaInfeccionesKM_TiemposTransplantes V2

14°CMayorm. nublado8:18 PM11/30/2025

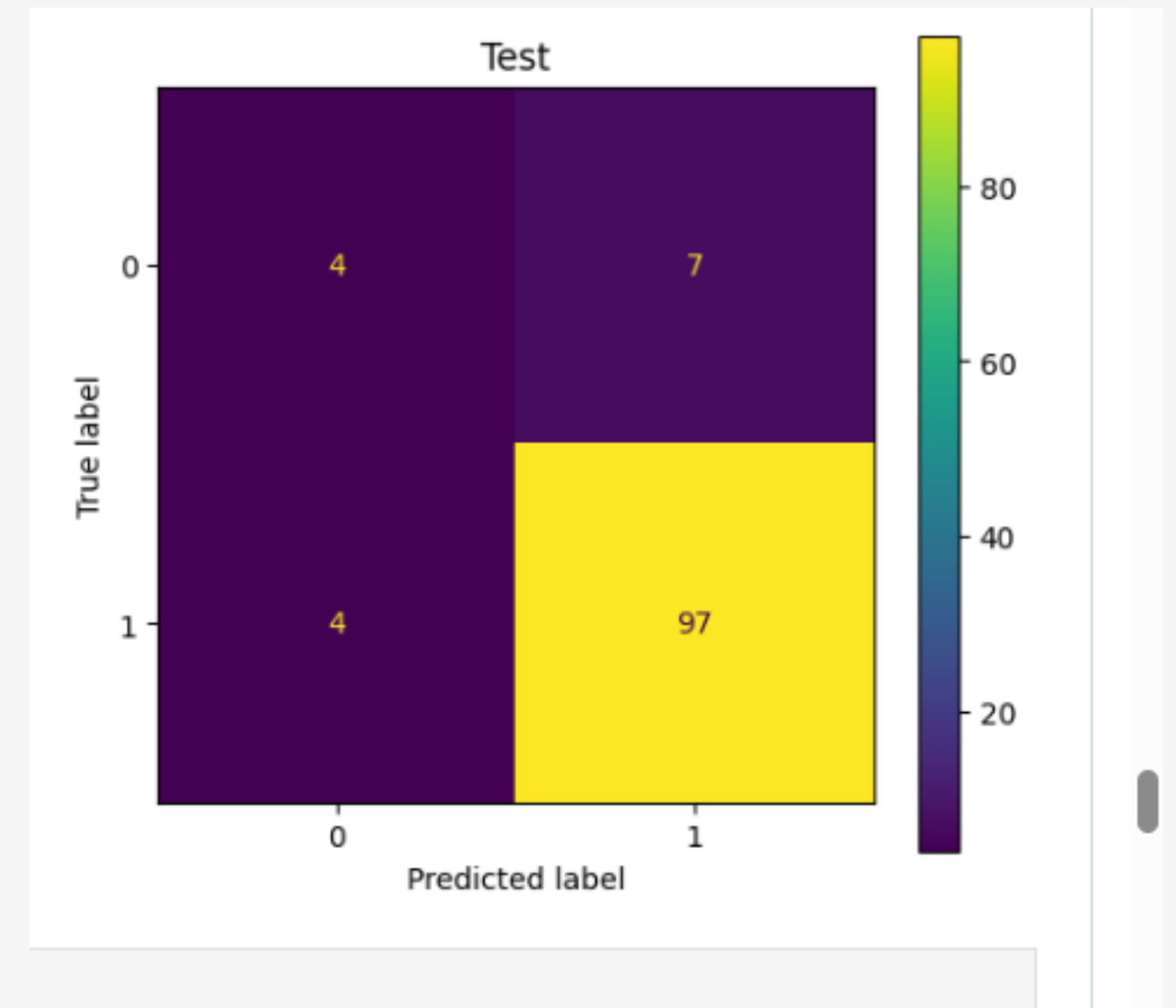


Modelos

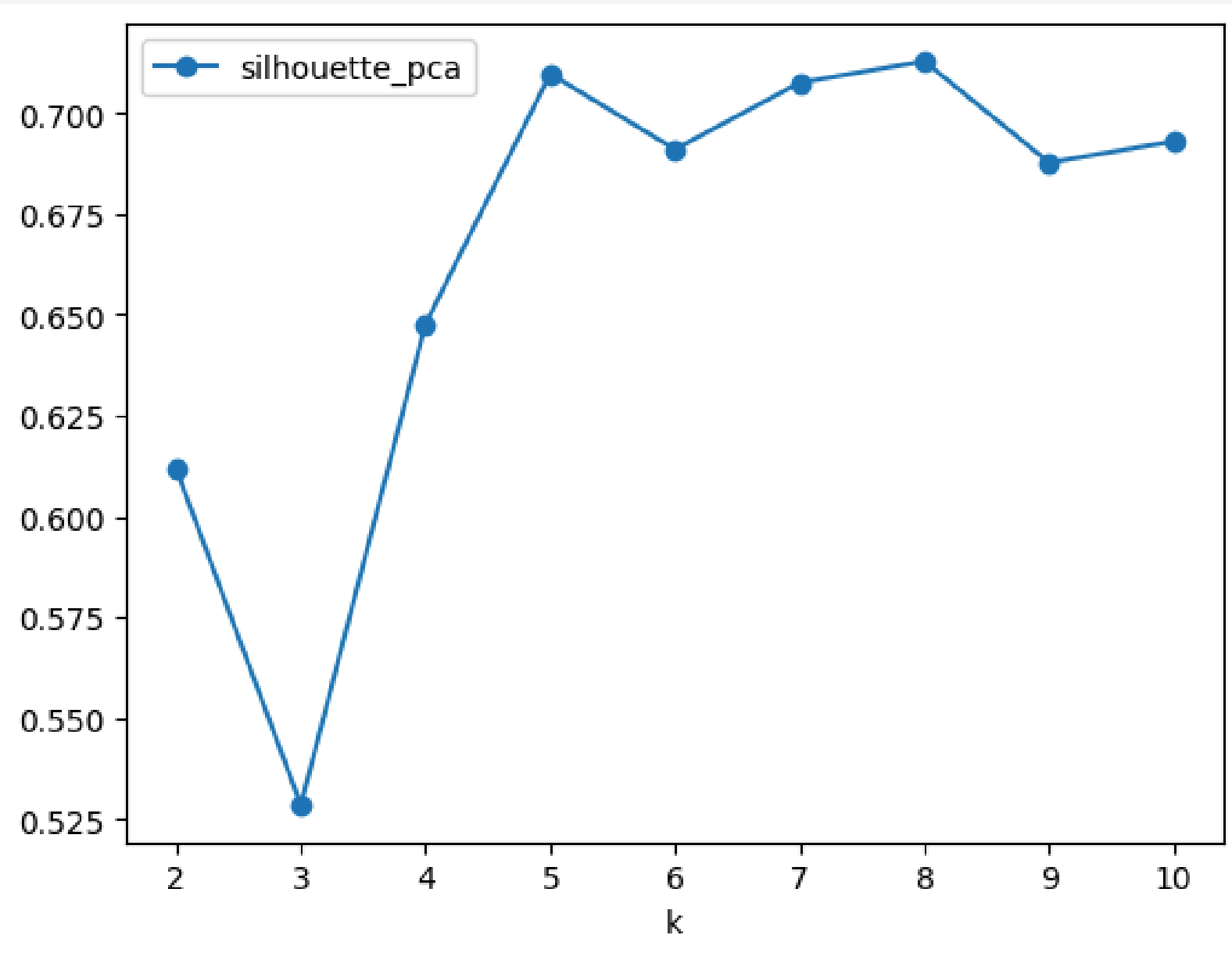
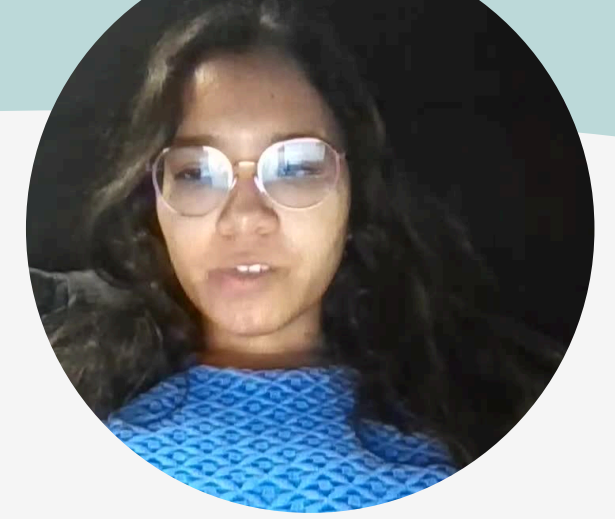
Modelo de clasificación

Regresión logística

- Variable a predecir: infección post trasplante
- Recall promedio de 0.9
- Variables relevantes: Diálisis, transfusión GRE y tiempo de estancia.

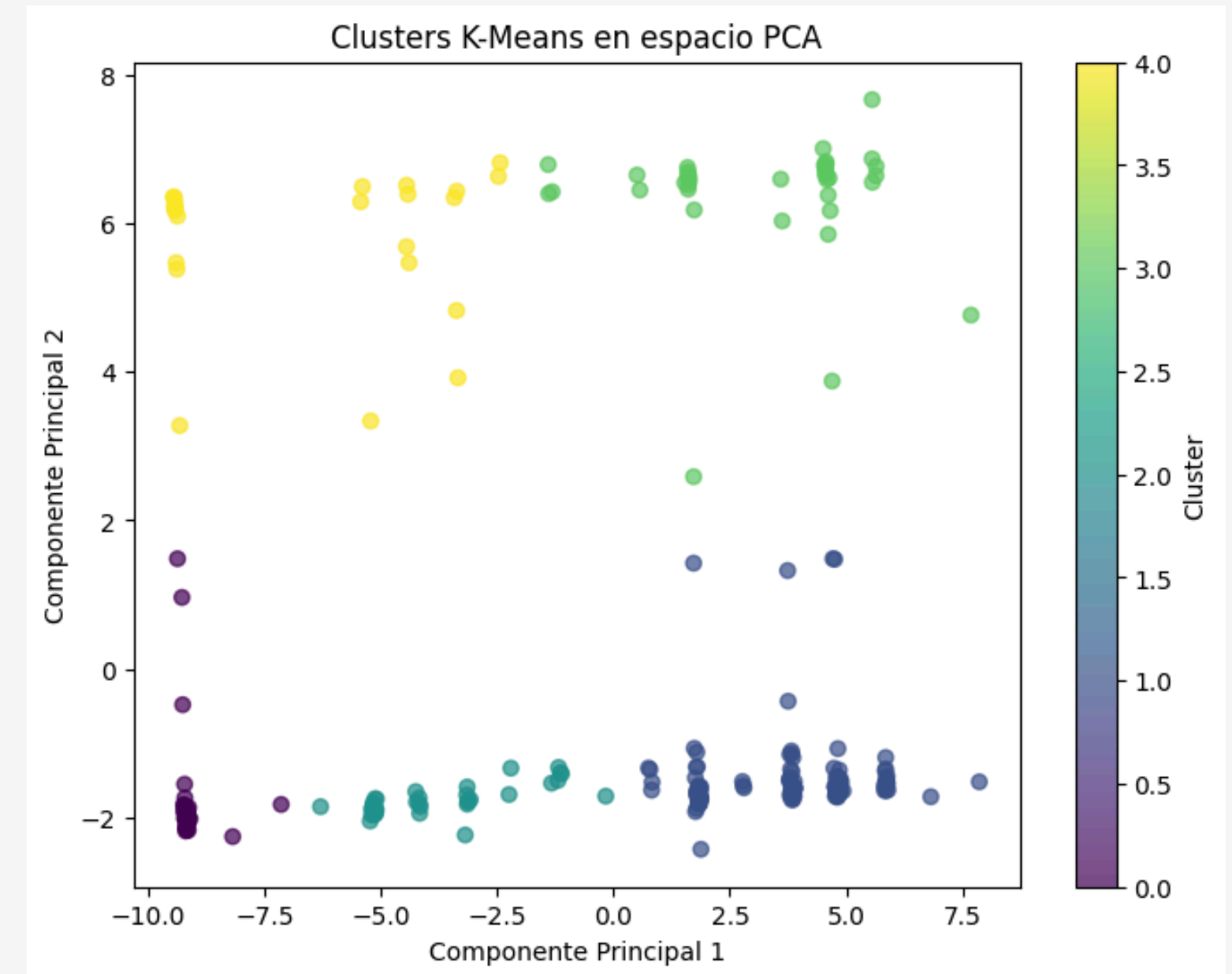


K-means Clustering

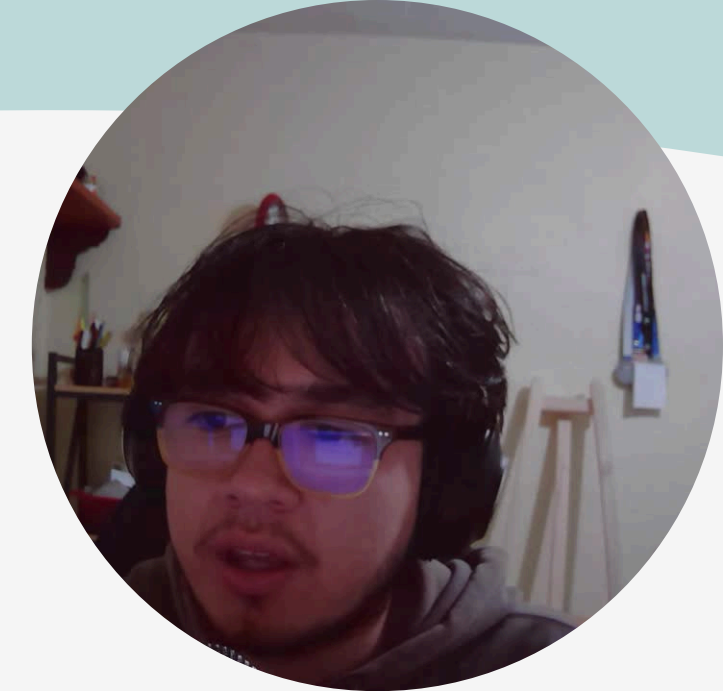


K = 5

Coeficiente de silueta = 0.71



**75.48% de la variabilidad
(PC1:55,70%, PC2: 19,78%)**



Conclusiones

- Se pasó de un Excel sin controles a un sistema estructurado, reduciendo errores y mejorando la calidad y confiabilidad de los datos.
- El dashboard desarrollado permite analizar métricas clínicas clave y apoyar decisiones que impactan directamente a los pacientes.
- La principal dificultad fue la mala calidad del dato y la ausencia de un diccionario estandarizado debido al registro manual.
- Para mejorar resultados futuros, la Fundación debe usar consistentemente la aplicación, permitiendo estandarizar datos y reentrenar modelos con mayor precisión.

¡Gracias!

