



**Fixing data**  
Data Driven Results Given

# Avance De Proyecto

Multivitaminico 60+

**SPRINT # 3**

Cierre del Proyecto



## Agenda de Hoy...

**01** Introducción

**02** Roles del Sprint

**03** Plan del Proyecto

**04** Repositorio Git Hub

**05** Pipeline - Cambios

**06** Modelo Clusters ML

**07** Deploy Streamlit

**08** Dash Board

**09** Conclusión







## Introducción

- Etapa de Data Analytics & MLOps
- Clusterización de Países para Recomendar países con mayor factibilidad de ser Rentables.
- Analisis de KPIs, contrastandolos con el modelo de Clusterización.
- Proceso Automatizado End-to-End



## Acerca de **Fixing Data**

Somos una joven y dinámica consultora de datos especializada en proyectos end-to-end. Nuestro equipo aborda proyectos que cubren todo el ciclo de vida de los datos. Con un enfoque multidisciplinario, ofrecemos soluciones integrales y valiosas para nuestros clientes.



**Ing. Fernando G. Cofone**  
PM & Data Engineer



**Ing. Paula Perosio**  
Data Analyst & Analytics



**Ing. Willian Jose Suarez**  
Data Analyst & Analytics



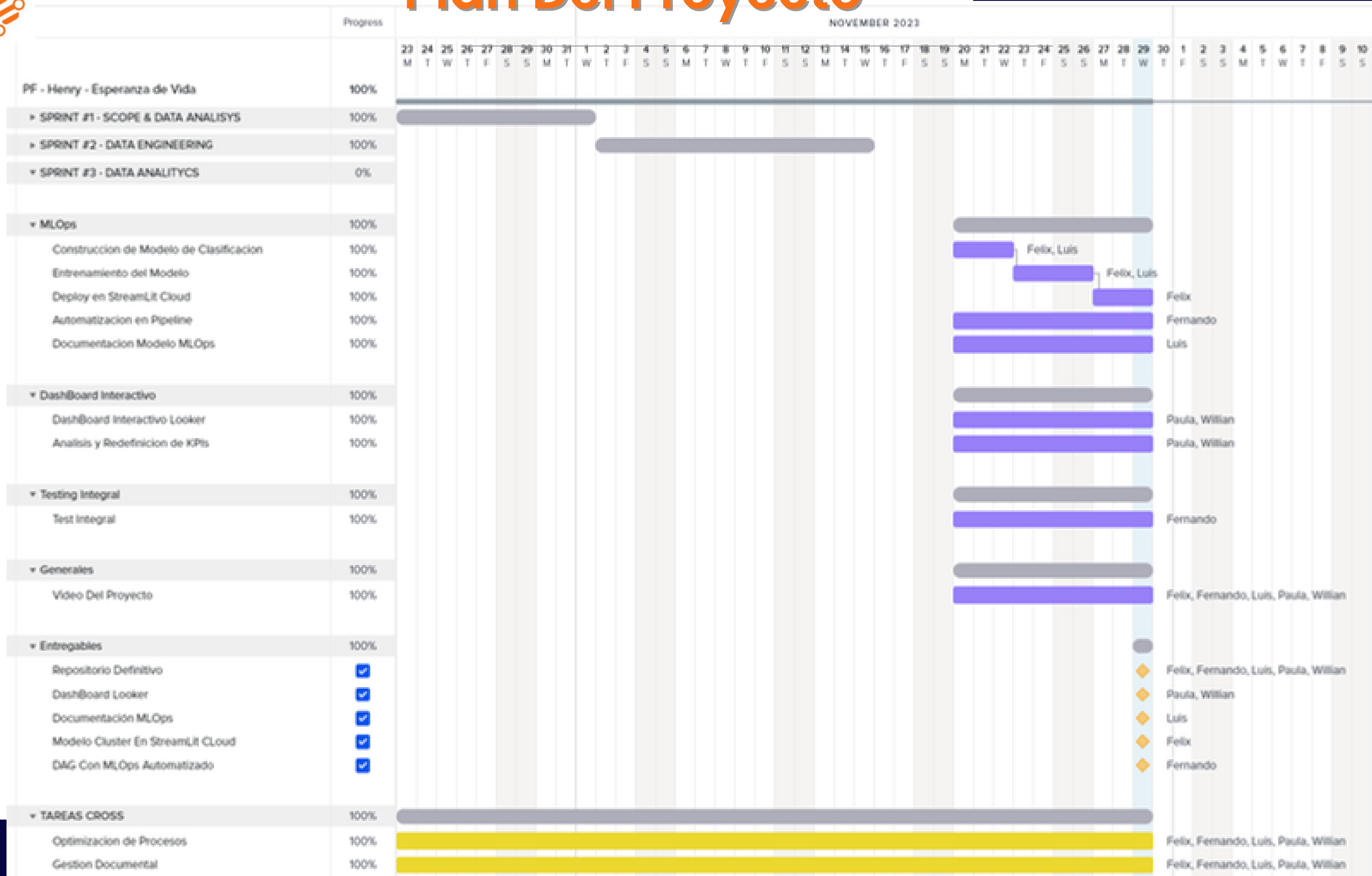
**Felix Contreras**  
MLOps Engineer



**Ing. Luis David Mendoza**  
MLOps Engineer



# Plan Del Proyecto





# REPOSITORIO GIT HUB



PF-Esperanza-De-Vida

Public

Pin

Unwatch 1

Fork 1

Star 0

main

8 branches

0 tags

Go to file

Add file

Code



fercof87

readme

1bcc328 13 hours ago 51 commits

Data-Analitics	Automatizacion MLOps y Documentacion	2 days ago
Data-API-BM	Renombrado de Folder	2 days ago
Data-Engineering	Pruebas integrales y funcion eliminar duplicados	13 hours ago
Documentacion	Pruebas integrales y funcion eliminar duplicados	13 hours ago
EDA	Cambios en Documentacion y Reordenamiento del Repositorio	17 hours ago
MLOps	ML-3	yesterday
Testing	Pruebas integrales y funcion eliminar duplicados	13 hours ago
src	readme	13 hours ago
.gitignore	Pruebas integrales y funcion eliminar duplicados	13 hours ago
README.md	readme	13 hours ago
requirements.txt	delete io	2 days ago

README.md

## Proyecto Final - Data Science - Henry BootCamp

### About

Proyecto Final - Henry Bootcamp -  
Esperanza de Vida Al Nacer

Readme

Activity

0 stars

1 watching

1 fork

### Releases

No releases published

[Create a new release](#)

### Packages

No packages published

[Publish your first package](#)

### Contributors 5

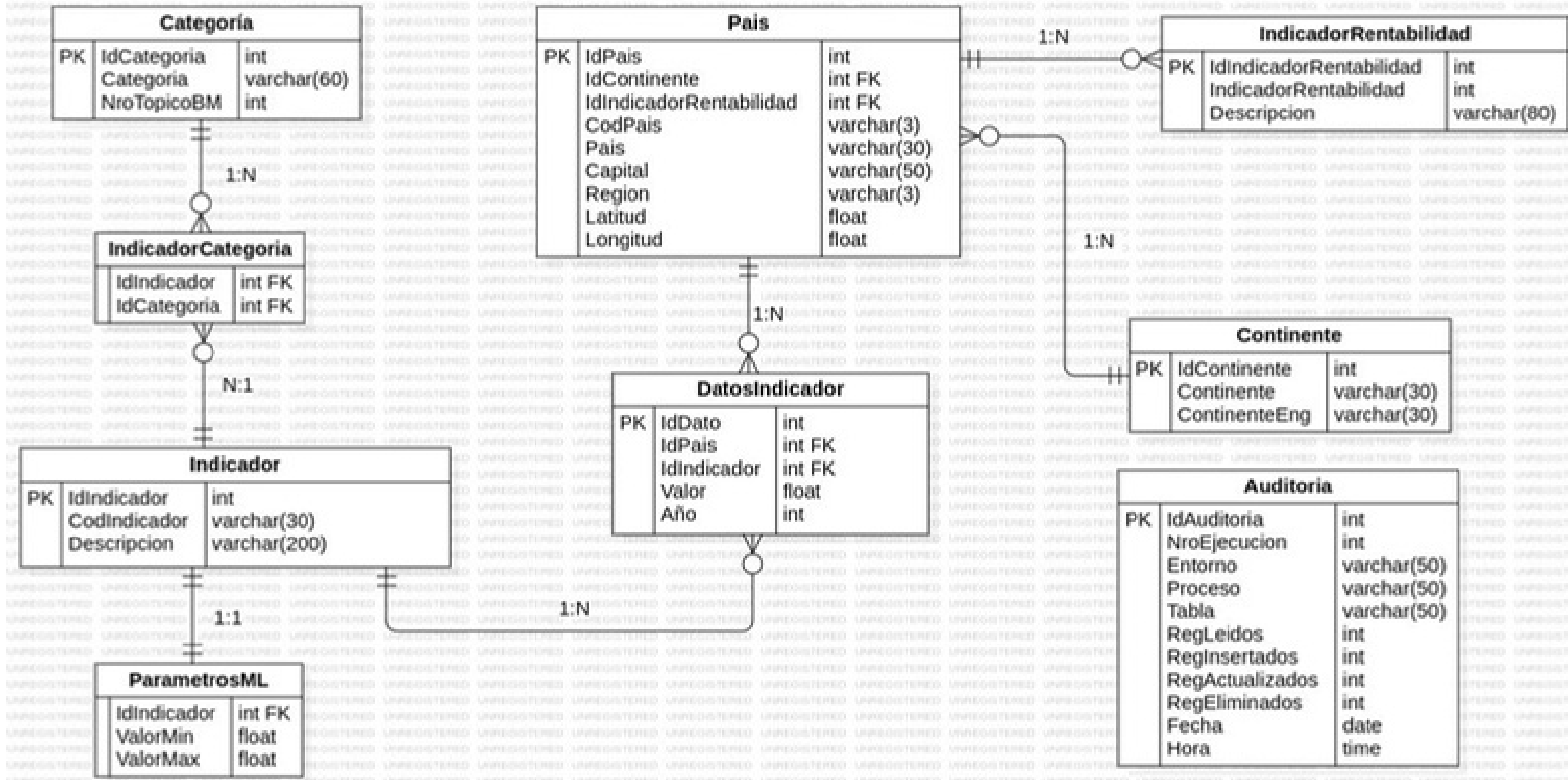


### Languages



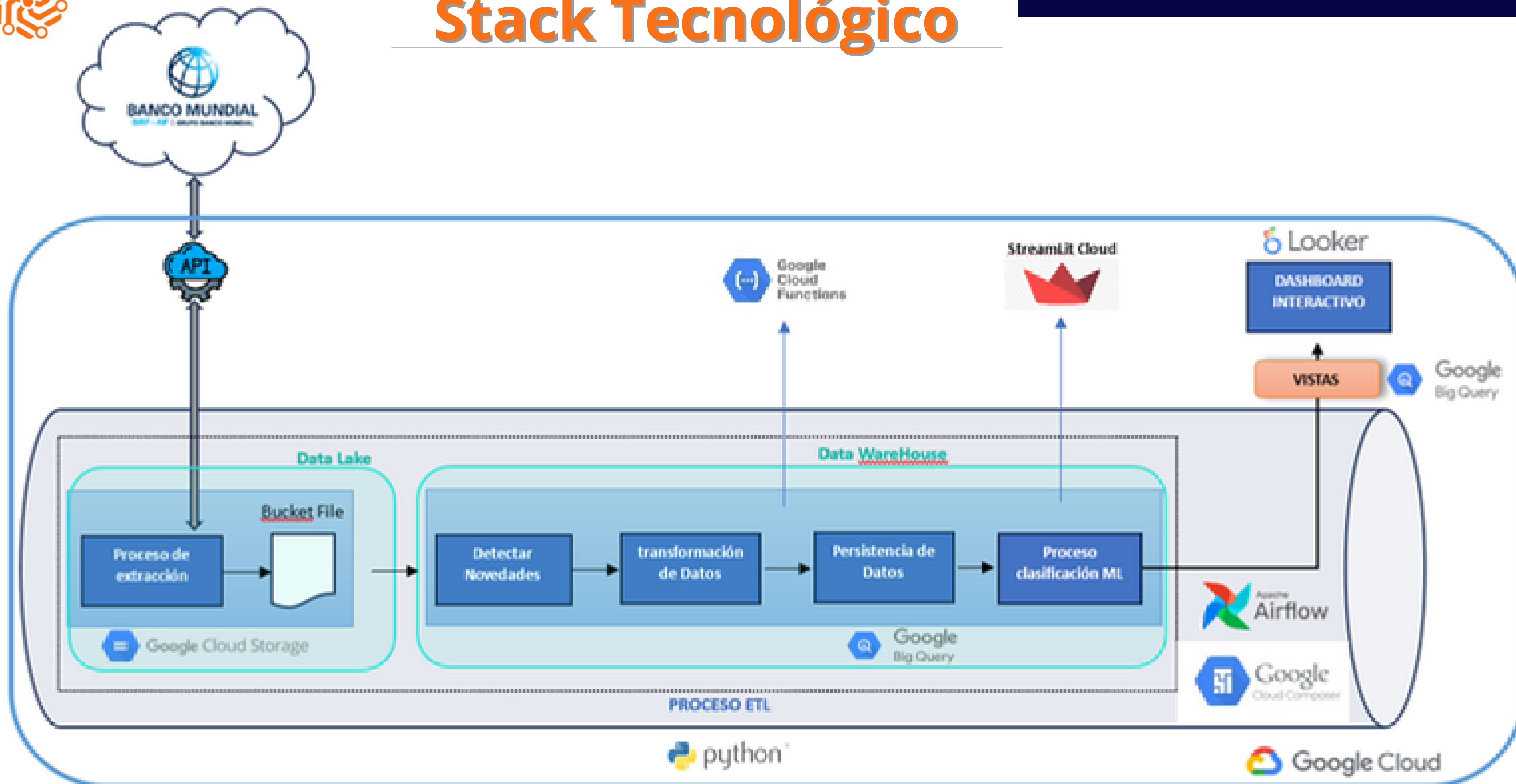
# Diagrama Entidad Relación

Diagrama Entidad Relación





# Stack Tecnológico







# DAG

Compositor

Entornos

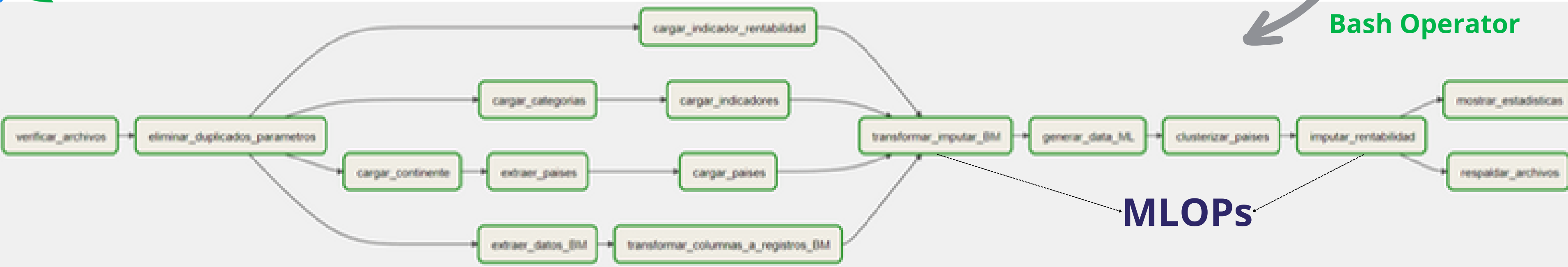
CREAR

BORRAR

Filtro

Filter environments

<div><input type="checkbox"/></div> Estado	Nombre <div>↑</div>	Ubicación	Versión de Composer	Versión de Airflow	Fecha y hora de creación	Hora de actualización	Webserver de Airflow	Lista de los DAG	Registros	Carpeta de DAG	Etiquetas
<div><input checked="" type="checkbox"/></div> <div></div>	<a href="#">composer.gcf.henry-esperanza</a>	us-central1	1.20.12	2.4.3	15/11/23, 09:16	15/11/23, 09:30	<a href="#">Airflow UI</a>	<div><div></div><div>DAG</div></div>	<div><div></div><div>Registros</div></div>	<div><div></div><div>DAG</div></div>	Ninguno





## Selección de indicadores



### **GDP per capita (current US\$)**

Producción promedio por individuo en el país.



### **Inflation, GDP deflator (annual %)**

Indica el cambio porcentual en el nivel general de precios en la economía.



### **Inflation, consumer prices (annual %)**

Cambio porcentual en el índice de precios al consumidor.



### **Life expectancy at birth, total (years)**

La esperanza de vida al nacer en años.



### **Population growth (annual %)**

Variación porcentual en la población en un año dado.



### **Ratio\_population ages 65 and above**

La proporción de la población total que tiene 65 años o más.



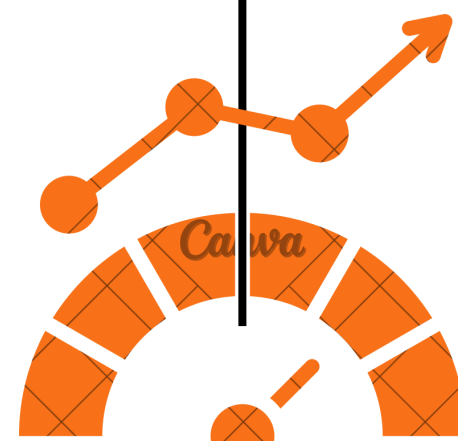
### **Ratio\_urban population**

La proporción de la población que vive en áreas urbanas



### **Urban population growth (annual %)**

La tasa de crecimiento anual de la población urbana.





# MODELO MACHINE LEARNING

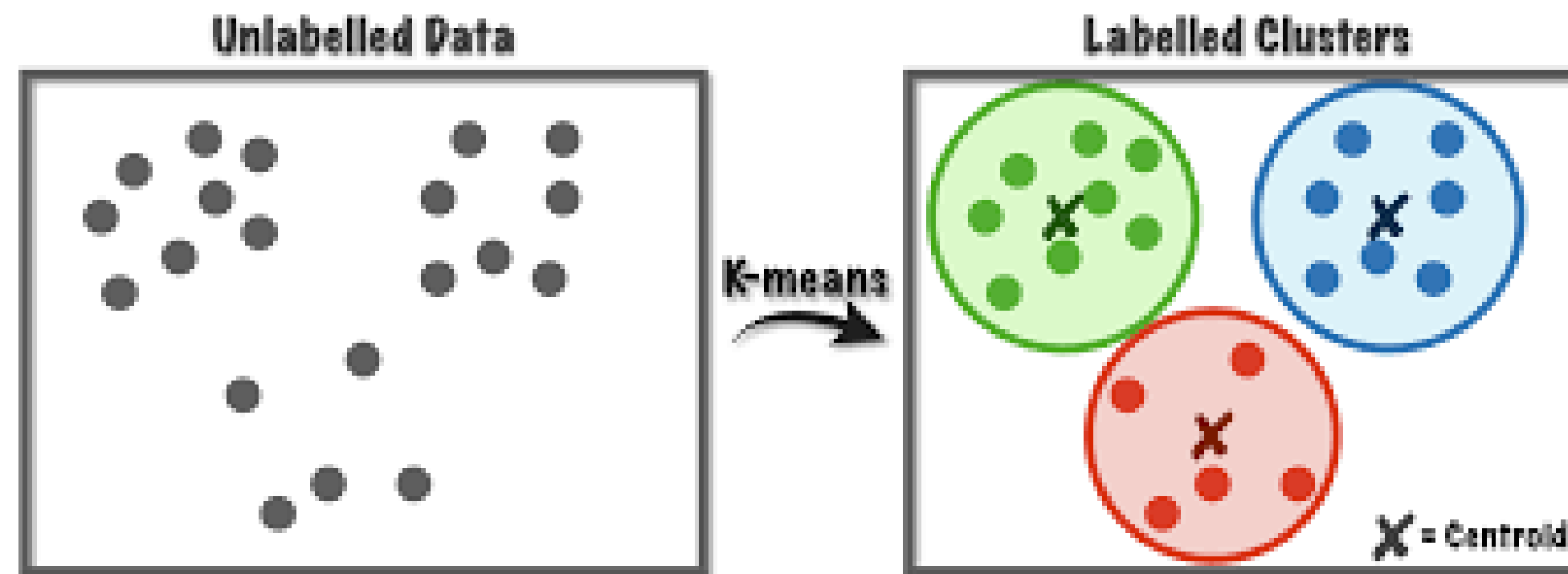
- Se ha decidido tomar como referencia el promedio de los valores registrados en los últimos 5 años (2017-2021).

Pais	GDP per capita (current US\$)	Inflation, GDP deflator (annual %)	Inflation, consumer prices (annual %)	Life expectancy at birth, total (years)	Population growth (annual %)	ratio_population ages 65 and above	ratio_urban population	Urban population growth (annual %)
AUS	55649.756424	2.702153	1.836326	82.929756	1.201004	0.159644	0.861286	1.331580
AUT	50282.610191	1.778456	1.951824	81.532683	0.495443	0.189736	0.585298	0.868419
BFA	797.775144	2.025153	1.148705	59.724600	2.735475	0.025614	0.299856	4.829883
BRA	8496.709453	6.049149	4.471526	74.406600	0.709181	0.090002	0.868184	1.003374
CAN	46751.935346	2.903484	1.985314	82.034541	1.139466	0.176072	0.814916	1.226117
CHE	85848.174807	0.122247	0.337790	83.612195	0.773234	0.185573	0.738636	0.842819
CHN	10378.507186	2.812982	1.993520	77.849600	0.350990	0.120393	0.602720	2.289982
COL	6231.102870	4.582955	3.419919	75.549000	1.570509	0.081996	0.810986	1.973864



# MODELO MACHINE LEARNING

- Se ha decidido el uso de un modelo no supervisado con el objetivo de segmentar los países en base a sus indicadores y poder recomendar los que se encuentren en una mejor situación.
- El algoritmo elegido es **K-means**.

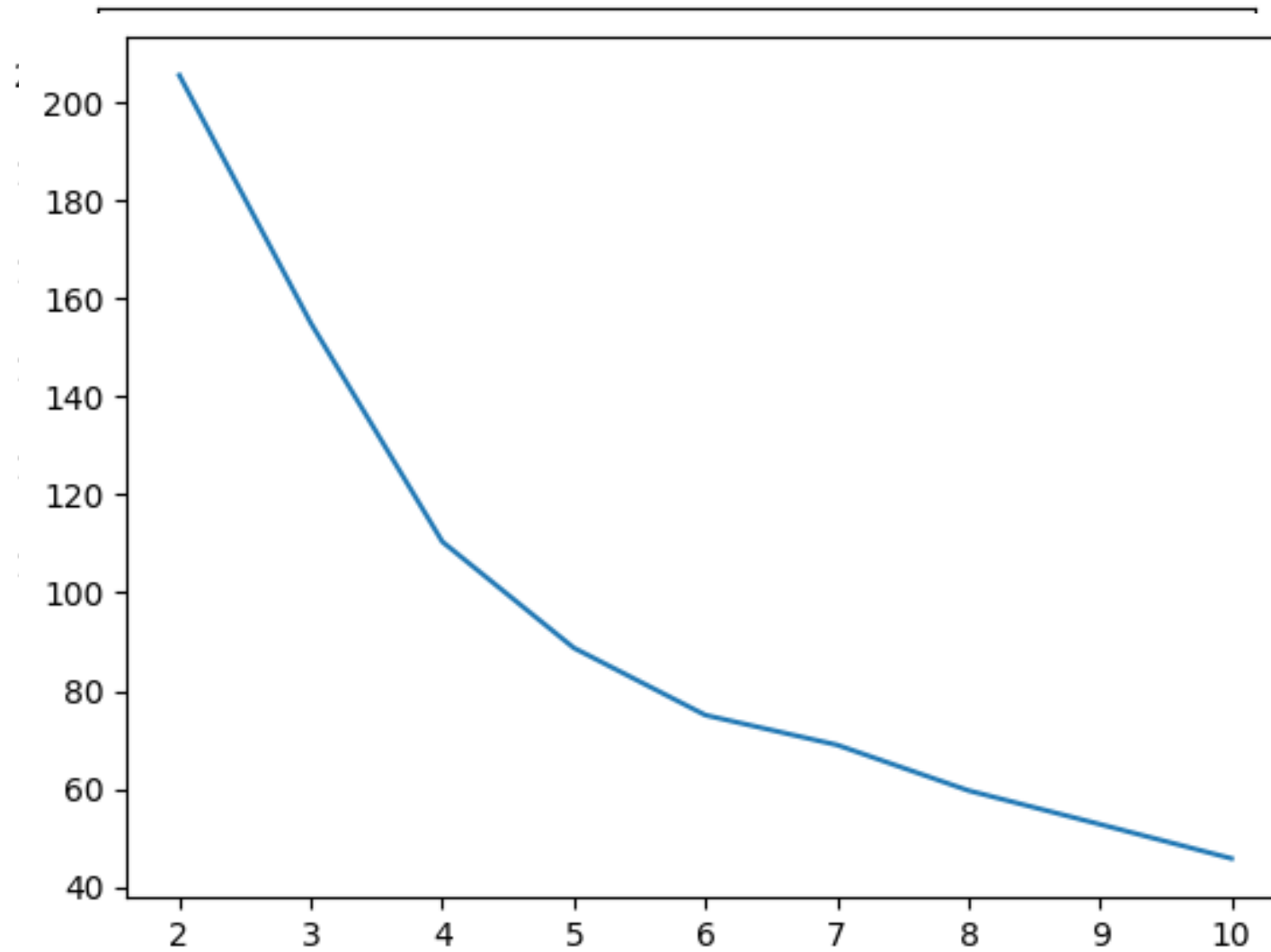




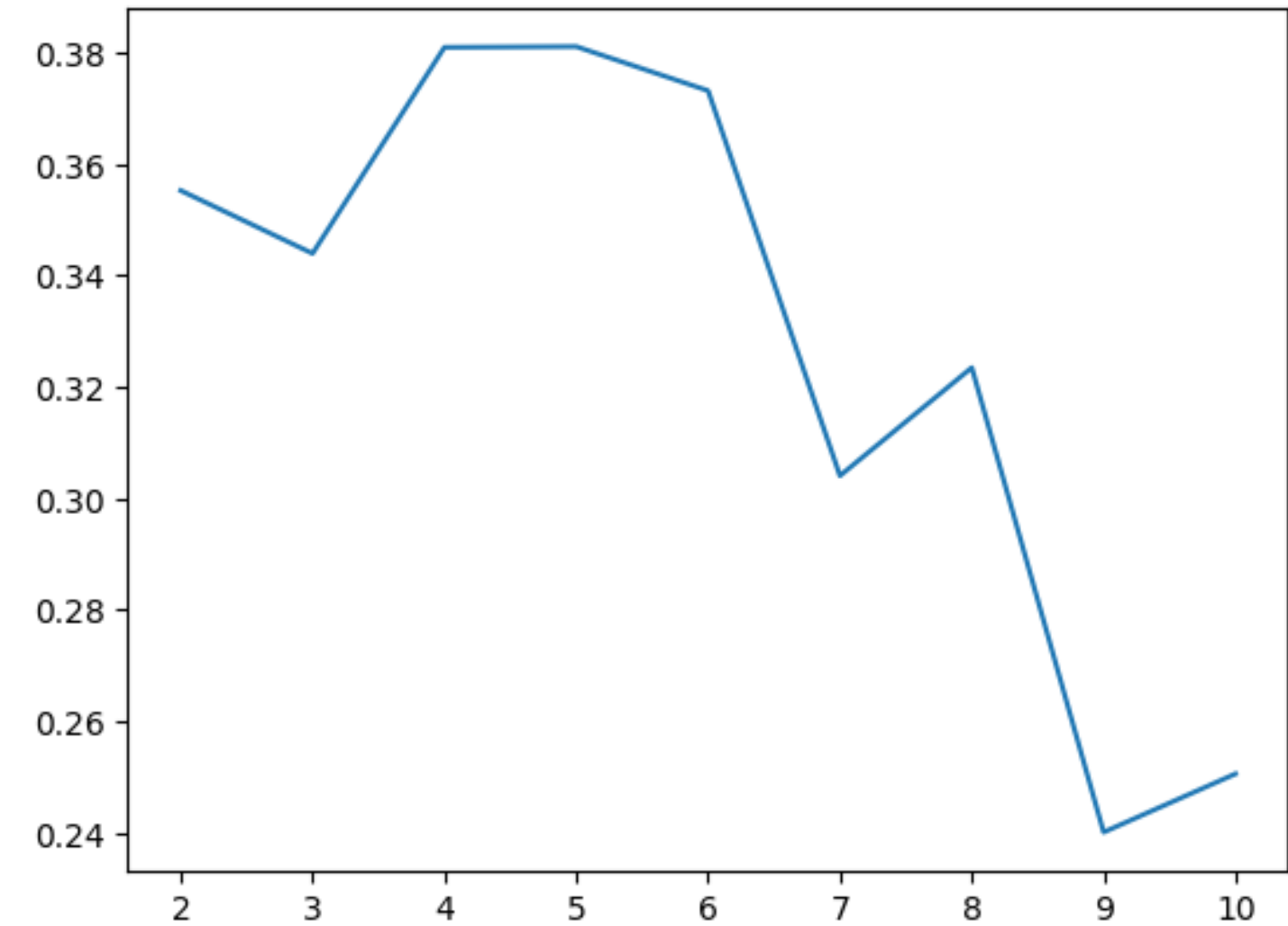


# MODELO MACHINE LEARNING

## Selección del número óptimo de clusters



**Método del codo**

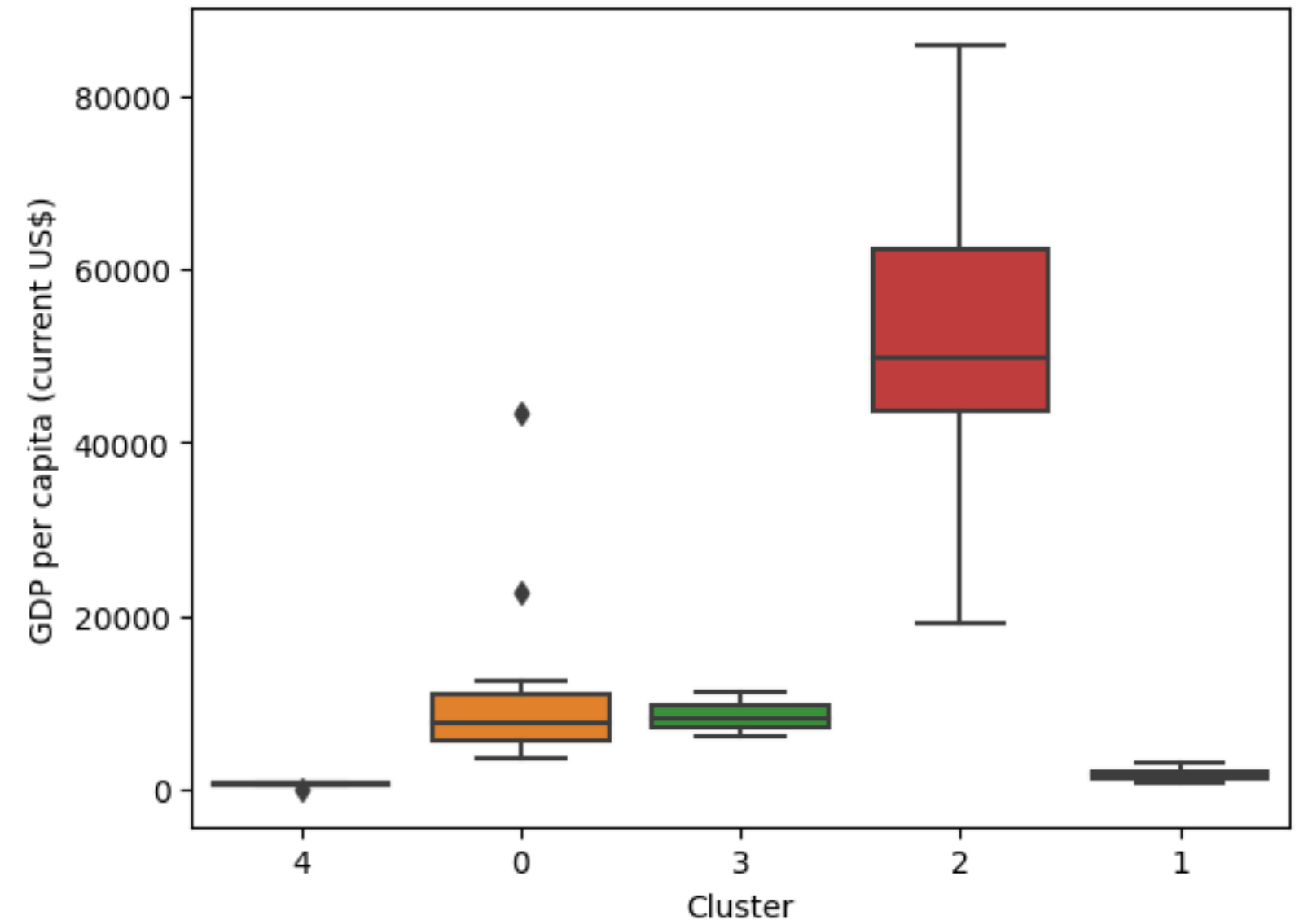
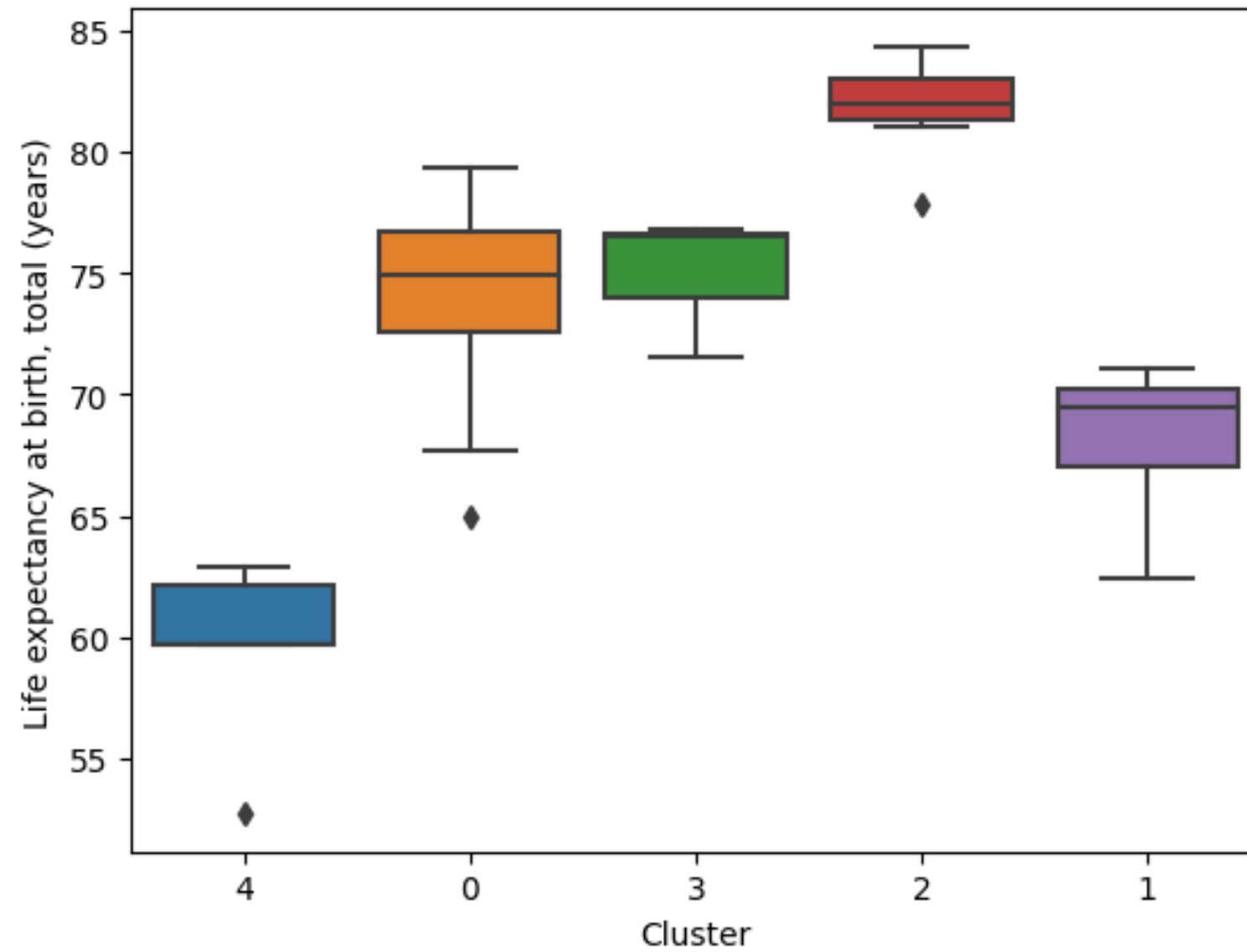


**Coeficiente de silueta**



# MODELO MACHINE LEARNING

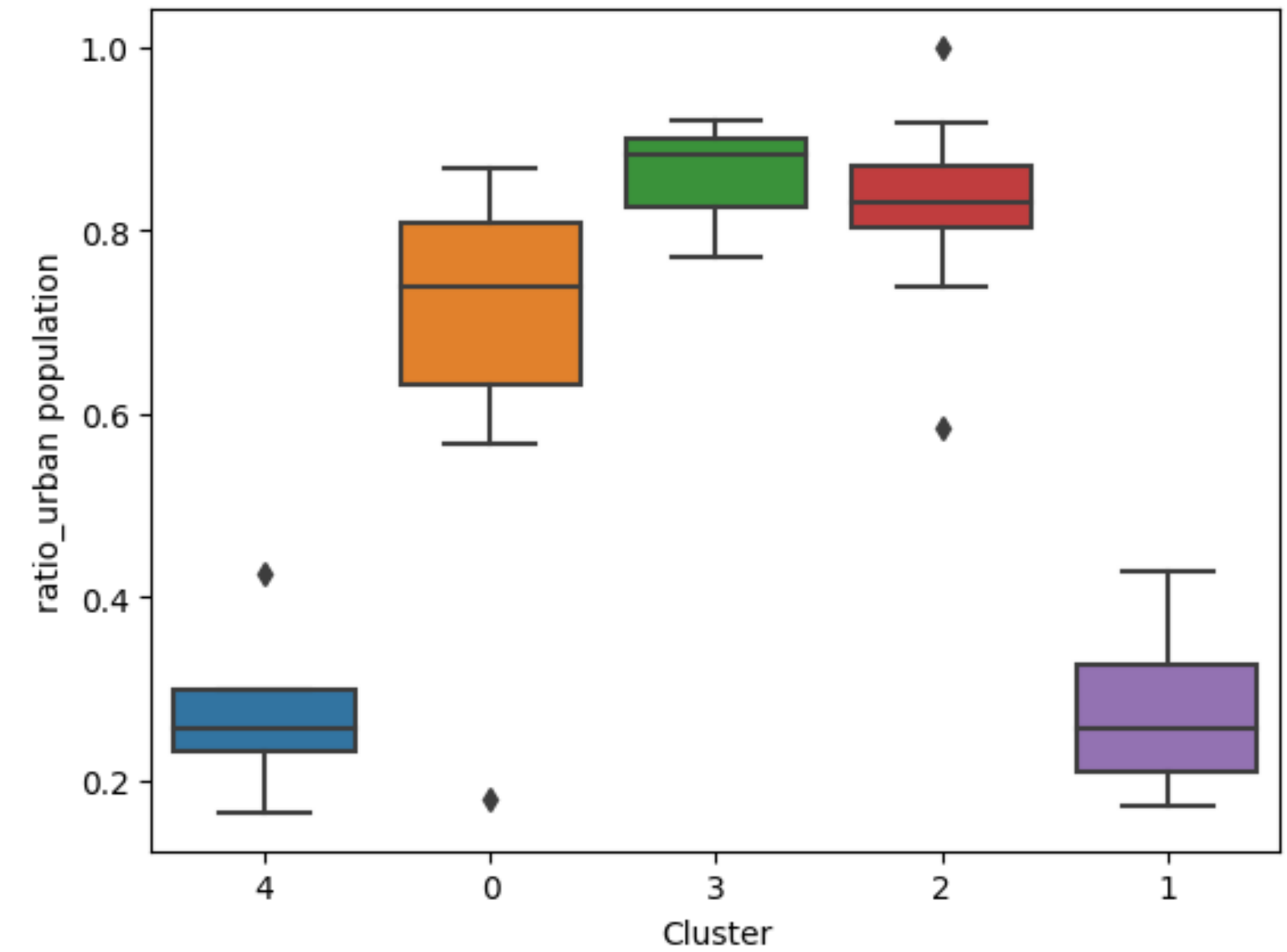
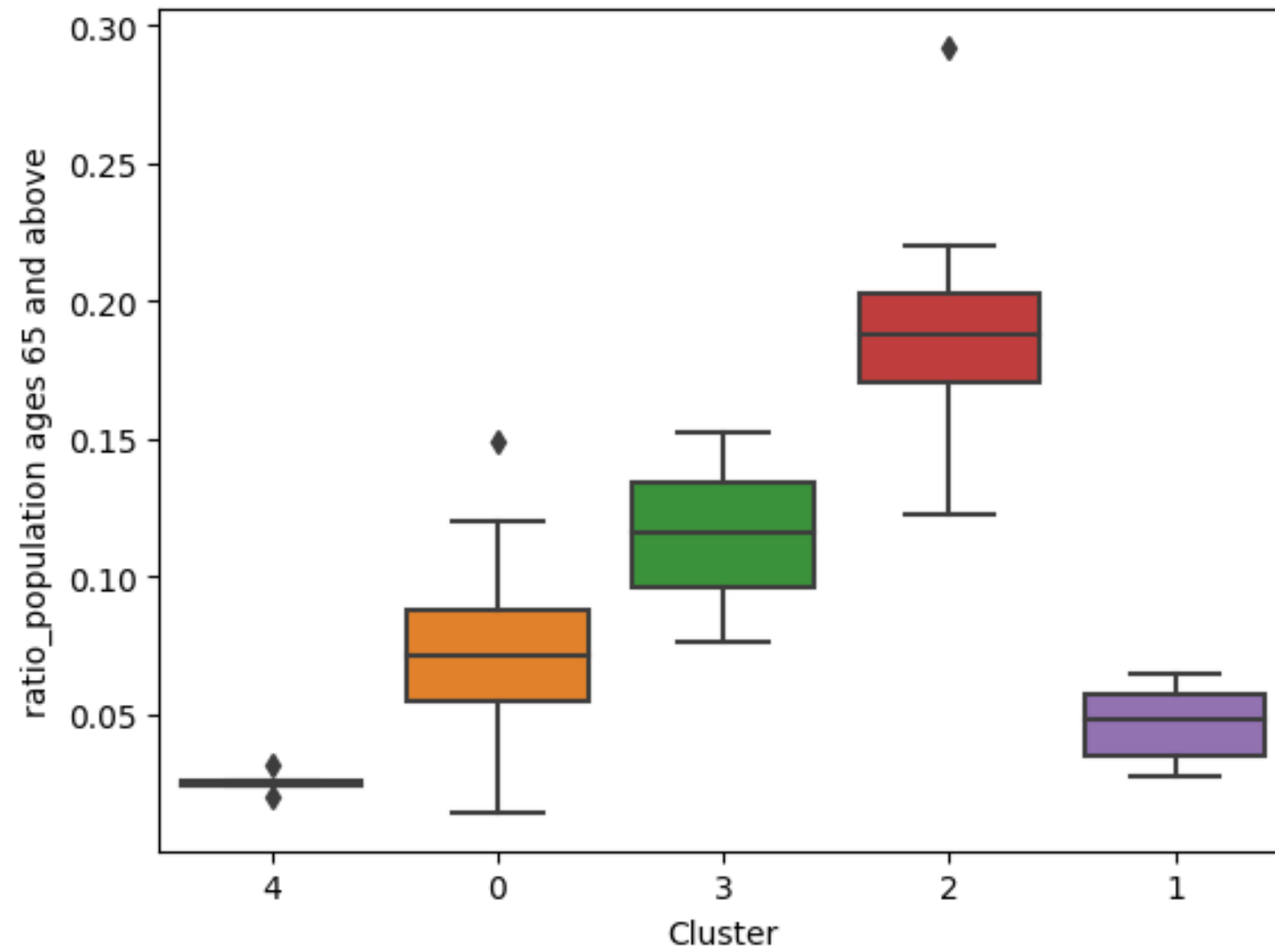
## Visualización de los clusters

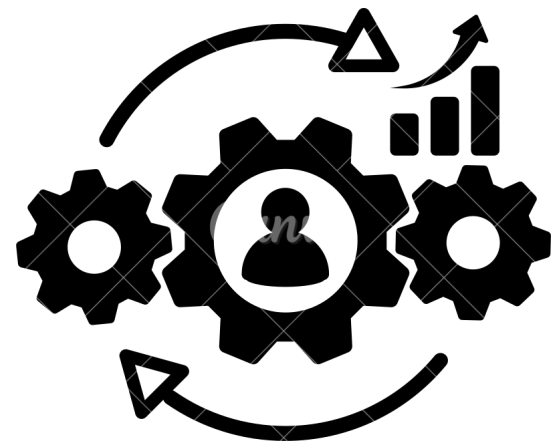




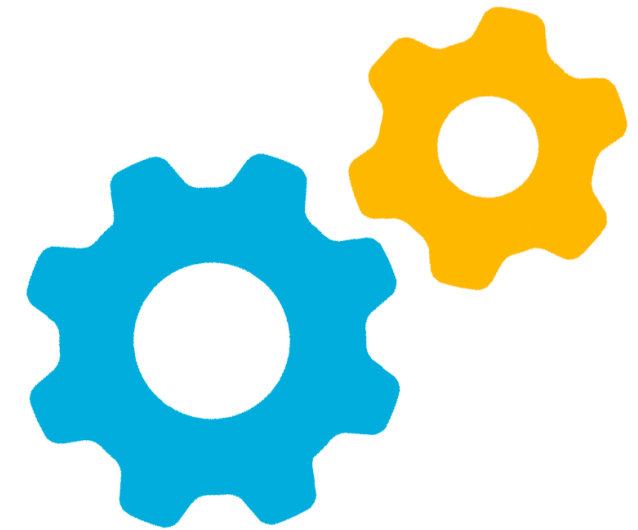
# MODELO MACHINE LEARNING

## Visualización de los clusters





- **Cuenta en Streamlit Cloud creada:**
  - Se ha creado una cuenta en Streamlit Cloud.
- **Vínculo establecido con el repositorio del proyecto:**
  - Se ha enlazado la cuenta de Streamlit Cloud con el repositorio del proyecto.



# Streamlit





- 1 - **Aumento** de la Esperanza de Vida mayor a 70 años en un 0,1%, periodo **quinquenal**
- 2 - **Aumento** de la Población Mayor a 65 años en un 2% **anual**
- 3 - **Aumento** de la Población Urbana en un 2% **anual**
- 4 - **Aumento** del Producto Interno Bruto (PBI) en un 0,25% **anual**
- 5 - **Mantener** la Inflación por debajo del 0%, periodo **quinquenal**

Key

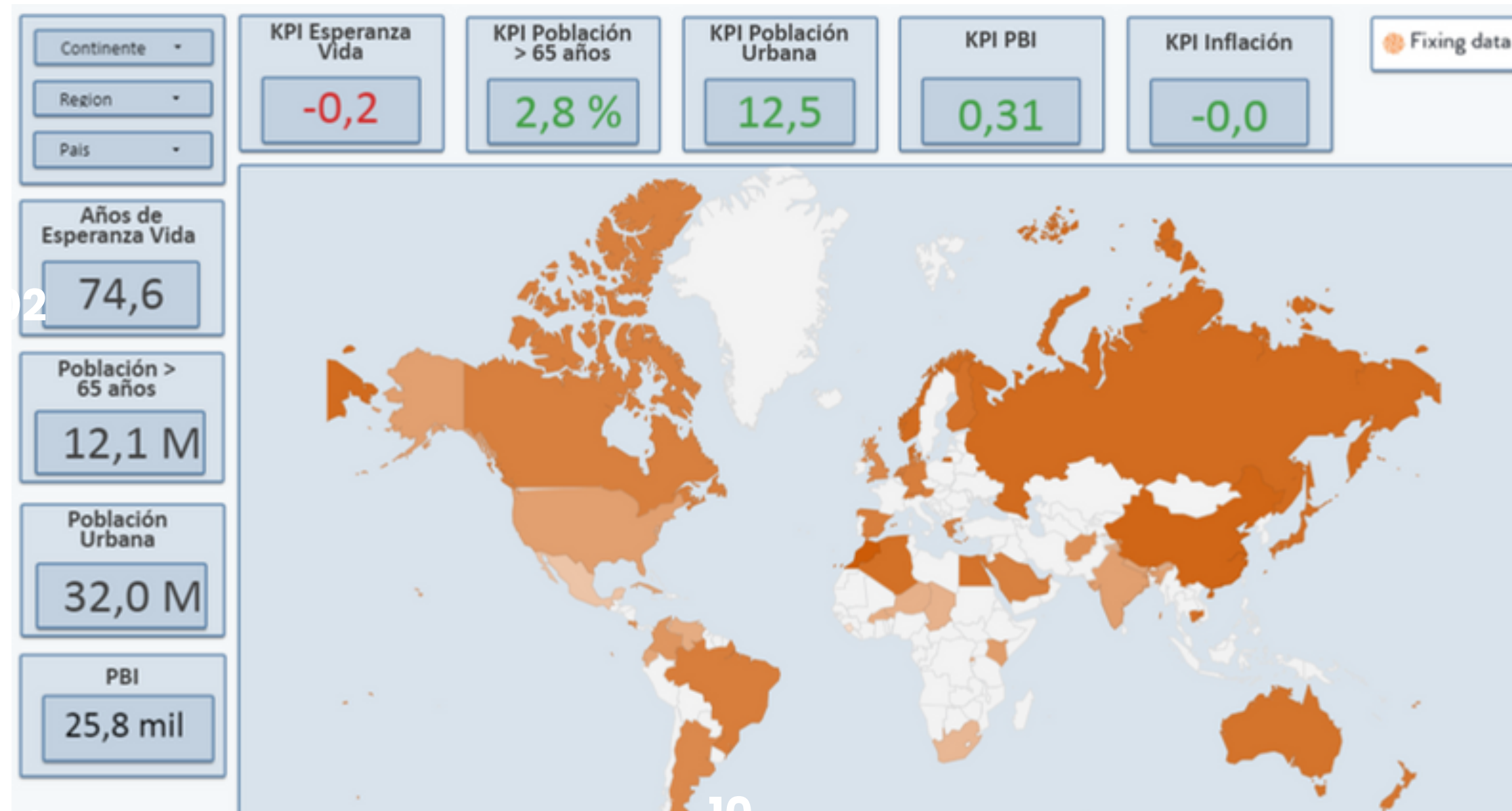
Performance

Indicator





# Dashboard Proyecto



Looker Studio



# Hitos del SPRINT # 3

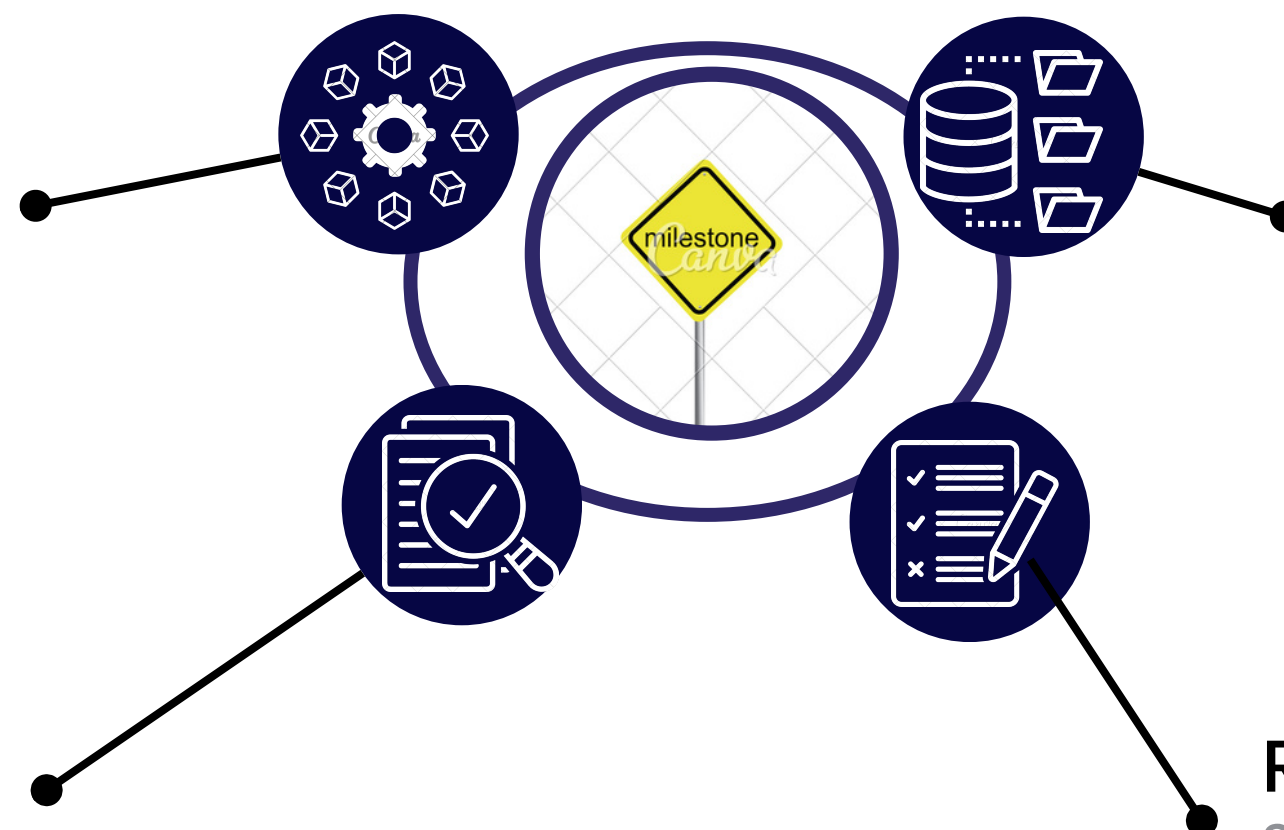
DATA ANALITYCS & MLOps

## Automatización de Pipeline con MLOps

Se automatizo el Modelo en el Pipeline y se hizo el Deploy en Stream Lit.

## Informe de Modelo Clusters

Informe resultante de la etapa de MLOps



## DashBoard Looker

Se construyo el tablero consumiendo los datos de BIG QUERY

## Repositorio Definitivo

Se disponibilizo el Repositorio en Git Hub





# Conclusión

- ✓ Finalización de la implementación de los requerimientos del Alcance Del Proyecto.
- ✓ Dashboard y Modelo de ML **Disponibilizados**
- ✓ Posibilidad de Realizar carga incremental para **disponer de mayor información.**
- ✓ Soporte en la Toma de decisiones para el **lanzamiento del nuevo Producto.**





# MUCHAS GRACIAS

Por Tu Atención!

THANK YOU  
Canva





**Fixing data**  
Data Driven Results Given

**Preguntas?**

