

ENTENDIMIENTO DE LA SITUACIÓN ACTUAL

La migración humana es un fenómeno que ha moldeado nuestras sociedades a lo largo de la historia y que sigue siendo relevante en el mundo actual. Las personas se desplazan por diversas razones, como la búsqueda de oportunidades económicas, el escape de situaciones de violencia o la adaptación a condiciones ambientales cambiantes. Estos factores, junto con el desarrollo tecnológico que ha facilitado la comunicación y el transporte, han generado una gran diversidad y complejidad en los flujos migratorios actuales.

En este contexto, el análisis de los datos migratorios se presenta como una herramienta esencial para comprender y abordar los desafíos y oportunidades asociados con la migración en la actualidad. Los datos recopilados sobre los movimientos de población, las rutas migratorias, las características demográficas y los factores impulsores proporcionan información valiosa que puede guiar la toma de decisiones informada en diversos niveles.

Esto no solo nos permitirá entender la magnitud y dirección de los movimientos de población, sino también explorar las posibles implicaciones socioeconómicas, culturales y políticas de estos flujos.

OBJETIVOS

- Analizar y comprender los flujos migratorios internacionales.
- Recolección y consolidación de datos históricos.
- Identificar países de origen y destino relevantes.
- Reconocer los principales factores para la migración.

ALCANCE

Análisis de Flujos Migratorios y Factores Motivadores: Recopilaremos y analizaremos datos históricos de flujos migratorios, centrándonos en los años comprendidos entre 2000 y 2020. Nos enfocaremos en los 10 países de origen y destino más frecuentes en este período. Exploraremos las razones detrás de estos movimientos, considerando factores socioeconómicos, políticos y medioambientales.

La elección del período de 2000 al 2020 se basa en la disponibilidad de datos completos y detallados para realizar un análisis sólido, mientras que la inclusión de años anteriores puede estar limitada por la falta de información completa en esos años debido a razones específicas de cada país.

OBJETIVOS Y KPIS ASOCIADOS (PLANTEO)

Los KPIs que hemos definido se basan en nuestra comprensión de los flujos migratorios y los factores que los impulsan, y están diseñados para generar un impacto positivo en la toma de decisiones y en la sociedad en general.

Documentación proyecto Flujos Migratorios.

- 1. Objetivo: Reducir la tasa de migración neta para los países con mayor flujo migratorio en 1% en el próximo año.**
 - **KPI:** Tasa de Migración Neta disminuida en un 1% en los países con los flujos migratorios más significativos.
- 2. Objetivo: Reducir la tasa de desempleo en el país de origen en un 3% en los próximos 2 años.**
 - **KPI:** Tasa de Desempleo en el país de origen reducida en un 3% en un período de 2 años.
- 3. Objetivo: Reducir la tasa de desempleo en el país de destino en un 3% en los próximos 2 años.**
 - **KPI:** Tasa de Desempleo en el país de destino reducida en un 3% en un período de 2 años.
- 4. Objetivo: Reducir en un 5% las muertes y desapariciones por migraciones irregulares.**
 - **KPI:** Reducción del 5% en el número de muertes y desapariciones relacionadas con migraciones irregulares.
- 5. Objetivo: Reducir el índice de conflictos en un 2% en los próximos 3 años.**
 - **KPI:** Disminución del 2% en el índice de conflictos en un período de 3 años.

SOLUCIÓN PROPUESTA

Nuestra solución se centrará en un enfoque estructurado y colaborativo que nos permitirá alcanzar los objetivos establecidos y crear un producto final integral y significativo. A continuación, detallamos las tareas clave que realizaremos para cumplir con nuestros objetivos y los productos que surgirán de cada etapa:

Metodología y Organización:

Adoptaremos una metodología ágil para el desarrollo de nuestro proyecto, utilizando Scrum como marco de trabajo. Organizaremos nuestro trabajo en sprints de una semana, con reuniones diarias de seguimiento y demos al final de cada sprint. Esto nos permitirá mantener una comunicación constante, abordar desafíos a medida que surjan y adaptarnos rápidamente a cambios.

Distribución de Tareas y Roles:

- **Product Owner:** representante ONG.
- **Scrum Master:** Mentor.
- **Equipo de Data Science:** Brenda Jaras, Kevin Bonin, Miller Rodríguez y Angie Arango.

Documentación proyecto Flujos Migratorios.

Roles y Responsabilidades del equipo:

Data Engineer: Brenda Jaras

- **Responsabilidad:** El Data Engineer se encargará de diseñar, construir y mantener la infraestructura de datos, asegurando la disponibilidad, calidad y procesamiento eficiente de los datos.

Tareas Clave:

- **Crear y mantener el Data Pipeline:** Diseñar, implementar y mantener un flujo de datos automatizado desde múltiples fuentes hasta el Data Warehouse.
- **Diseñar y Gestionar la Base de Datos y Data Warehouse:** Crear y mantener una estructura de base de datos óptima para el almacenamiento y análisis de datos migratorios.
- **Automatizar Procesos de ETL:** Utilizar herramientas como AWS Glue para automatizar el procesamiento y la carga de datos en el Data Warehouse.

Data Scientist: Kevin Bonin y Miller Rodriguez

- **Responsabilidad:** El Data Scientist se enfocará en analizar los datos, identificar patrones y tendencias, y desarrollar modelos predictivos para comprender mejor los flujos migratorios.

Tareas Clave:

- **Análisis Exploratorio de Datos (EDA):** Realizar análisis detallados para descubrir relaciones, patrones y características significativas en los datos migratorios.
- **Desarrollo de Modelos Predictivos:** Utilizar técnicas de Machine Learning para construir modelos que pronostiquen patrones futuros en los flujos migratorios en base a factores históricos y contextuales.
- **Definir KPIs Relevantes:** Identificar y definir los KPIs que medirán el impacto de los flujos migratorios y proporcionarán información para la toma de decisiones.

Data Analyst: Angie Arango

- **Responsabilidad:** El Data Analyst se encargará de traducir los datos en información comprensible y presentable, destacando los aspectos clave para la toma de decisiones.

Tareas Clave:

- **Creación de Dashboards Interactivos:** Diseñar y desarrollar un dashboard interactivo que muestre visualmente los datos migratorios relevantes y los KPIs identificados.
- **Generación de Informes:** Preparar informes y presentaciones que comuniquen los resultados y hallazgos de manera clara y concisa.

Documentación proyecto Flujos Migratorios.

- Colaboración Interdisciplinaria: Trabajar en conjunto con el Data Engineer y el Data Scientist para asegurar la correcta interpretación y visualización de los datos.

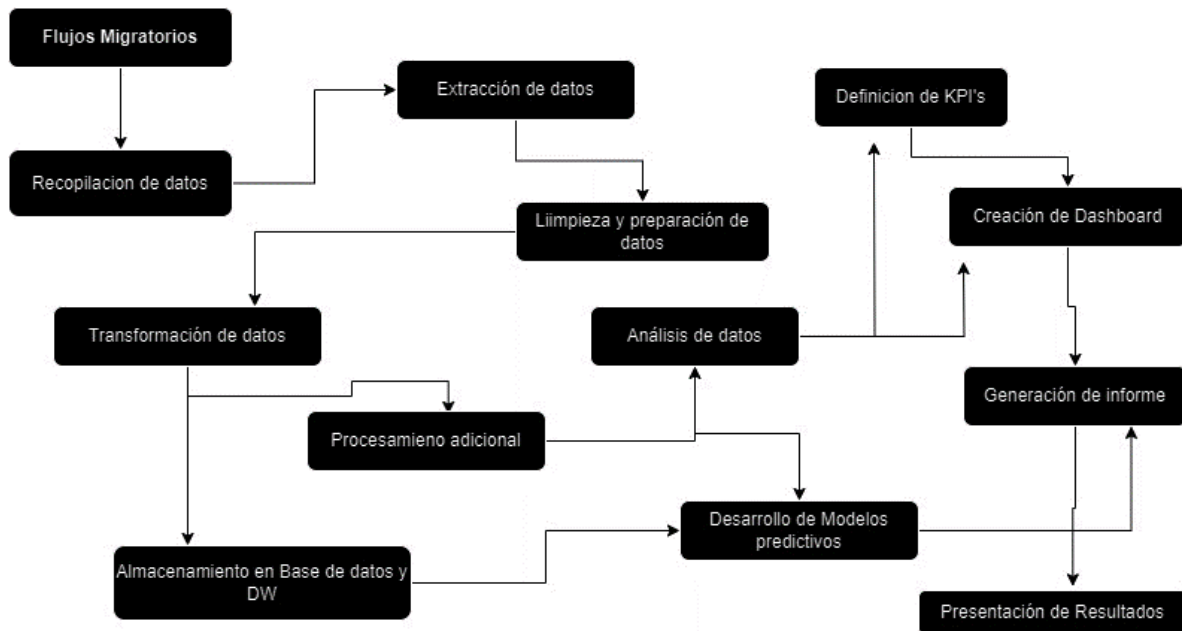
Colaboración y Comunicación:

Los tres roles trabajarán en estrecha colaboración, compartiendo sus hallazgos y contribuciones en reuniones regulares de equipo.

→ En el primer sprint, nos enfocaremos en la puesta en marcha del proyecto y en trabajar con los datos. Definiremos cuatro KPIs que medirán el impacto de los flujos migratorios, documentaremos el alcance del proyecto y realizaremos un análisis exploratorio de los datos. Además, estableceremos un repositorio en GitHub, implementaremos el stack tecnológico elegido y estableceremos la metodología de trabajo. Al finalizar este sprint, tendremos un flujo de trabajo establecido y una base sólida para el análisis de datos.

FLUJO DE TRABAJO

En esta sección, presentamos el diseño detallado del flujo de trabajo que seguiremos en nuestro proyecto, abarcando desde la recopilación de datos hasta la presentación de resultados. Cada etapa es crucial para alcanzar nuestros objetivos y crear un análisis completo de los flujos migratorios.



STACK TECNOLÓGICO:

- **Lenguaje de Programación: Python**

Utilizaremos Python como lenguaje principal para el desarrollo de scripts y código necesario en todas las etapas del proyecto. Python es ampliamente conocido en la comunidad de análisis de datos y es compatible con una variedad de bibliotecas y frameworks para análisis y modelado.

- **Extracción y Limpieza de Datos:**

- * **Bibliotecas: pandas, NumPy**

Utilizaremos la biblioteca pandas para la manipulación y limpieza de datos. pandas nos permitirá cargar, transformar y limpiar los datos de manera eficiente. NumPy se usará para operaciones numéricas y cálculos matriciales.

- **Visualización de Datos:**

- * **Bibliotecas: Matplotlib, Seaborn, geopandas**

Utilizaremos Matplotlib y Seaborn para crear visualizaciones claras y efectivas de los datos. geopandas nos permitirá trabajar con datos geoespaciales.

- **Base de Datos:**

- * **Base de Datos Relacional: MySQL**

Utilizaremos MySQL como sistema de gestión de base de datos relacional para almacenar datos estructurados.

- **Ingesta de Datos:**

- * **Servicio: AWS Lambda**

Utilizaremos AWS Lambda para la ingesta de datos, permitiendo el procesamiento de datos en tiempo real.

- **Almacenamiento en Data Lake:**

- * **Servicio: AWS S3**

Utilizaremos AWS S3 como Data Lake para almacenar y gestionar grandes volúmenes de datos de manera escalable y segura.

- **Transformación de Datos:**

- * **Servicios: AWS Lambda y AWS Glue**

Utilizaremos AWS Lambda para transformaciones de datos específicas y AWS Glue para procesos de ETL más complejos y automatizados.

- **Colaboración y Control de Versiones:**

- * **Plataforma: GitHub**

Documentación proyecto Flujos Migratorios.

Utilizaremos GitHub para la colaboración en equipo, el control de versiones del código y la gestión de recursos del proyecto.

- **Creación de Dashboards Interactivos:**

- * **Herramienta: Power BI**

Utilizaremos Power BI para diseñar y desarrollar dashboards interactivos que visualicen de manera efectiva los resultados del análisis de datos.

Este stack tecnológico integral permitirá llevar a cabo las diferentes etapas del proyecto, desde la recopilación y transformación de datos hasta la visualización de resultados, siguiendo una metodología ágil y colaborativa.

ANÁLISIS SOBRE LOS DATOS

Entendemos la importancia de trabajar con datos de alta calidad y, por lo tanto, realizaremos un análisis detallado de los metadatos de los datos con los que vamos a trabajar en nuestro proyecto de análisis de flujos migratorios internacionales. A continuación, proporcionamos una descripción de los metadatos relevantes:

Fuente de Datos:

Los datos se obtendrán de fuentes confiables, que incluyen organismos internacionales como la Organización de las Naciones Unidas (ONU), la Organización Internacional para las Migraciones (OIM), el Banco Mundial, entre otras.

Confiabilidad de las Fuentes:

Las fuentes utilizadas son ampliamente reconocidas y utilizadas en estudios de migración. Son confiables y proporcionan datos actualizados y precisos sobre flujos migratorios a nivel global.

Descripción de los Datos:

Los datos contendrán información sobre flujos migratorios internacionales, incluyendo:

- País de origen de los migrantes.
- País de destino de los migrantes.
- Porcentaje de mujeres, en país destino, y país origen.
- Factores de migración: muertes por conflictos, esperanza de vida, nivel de corrupción, inflación del deflactor del PIB, tasa de desempleo, crecimiento del PIB, homicidios intencionales, emisiones de CO2, migración neta y población neta.
- Información geográfica nombre del País, Código del país.
- Brecha de tiempo de análisis, Año, entre el 2000 y 2020.
- Población de refugiados por país de origen.

Documentación proyecto Flujos Migratorios.

- Población de refugiados por país de asilo.

Tipos de Datos:

Los tipos de datos incluirán números enteros para datos demográficos, números decimales para indicadores económicos, porcentajes, texto para descripciones y categorías

Método de Adquisición:

Los datos se adquirirán a través de consultas a bases de datos en línea, descargas de informes en formato csv y documentos de organismos internacionales, y acceso a datos gubernamentales en línea por medio de APIs.