



Flujos Migratorios en Argentina

La migración es un fenómeno global que afecta a muchas personas en todo el mundo. Hay muchas razones por las que la gente decide emigrar, incluyendo motivos económicos, políticos, sociales y de seguridad. Algunos de los desafíos más comunes que enfrentan los migrantes incluyen la discriminación, la falta de recursos, apoyo, y la separación de sus seres queridos.

El problema de la migración también tiene un impacto significativo en las comunidades de origen y destino. Por un lado, la emigración masiva puede socavar la economía y la estabilidad política de los países de origen. Además, los países de destino pueden verse abrumados por la demanda de recursos y servicios por parte de los migrantes, lo que puede llevar a tensiones entre los migrantes y la población local.

Para un argentino es posible encontrar mejores oportunidades de trabajo y salarios más altos que en Argentina, lo que puede mejorar la calidad de vida de una persona y proporcionar un futuro más estable. De igual manera, al tener acceso a un trabajo y un salario mejor, y a un costo de vida más bajo, puede ser posible mejorar la situación financiera de una persona y lograr sus metas y sueños.

En nuestro proyecto, queremos brindar a las personas, que viven en argentina o que deseen llegar a este país, de orientarlo cuál es el mejor lugar de destino. Para esto, puede ser útil recopilar y analizar datos de diferentes fuentes, como estadísticas gubernamentales, informes de organizaciones no gubernamentales y encuestas de opinión pública. Estos datos pueden utilizarse para calificar y comparar diferentes destinos en términos de los factores.

Además, es importante tener en cuenta las preferencias y necesidades individuales de cada emigrante, incluyendo sus motivos para emigrar, sus habilidades y antecedentes profesionales, sus redes de apoyo. La herramienta podría tener un



cuestionario o un sistema de recomendación que permita a los usuarios identificar sus prioridades y recibir recomendaciones personalizadas. Incluir información sobre el proceso de inmigración y el costo de vida en diferentes destinos, así como información sobre cómo establecerse en un nuevo país, cómo encontrar trabajo y vivienda.

Objetivo general

El objetivo general del proyecto es proporcionar una herramienta útil y accesible para ayudar a los emigrantes argentinos, y de igual manera a las personas que desean emigrar a Argentina. A tomar decisiones informadas sobre su destino de emigración. La herramienta se enfocará en datos objetivos y personalizados para cada usuario, y sería un recurso valioso para aquellos que buscan un futuro mejor y más estable.

Objetivo específico

1. Recopilar y analizar datos relevantes sobre diferentes destinos de emigración para calificar y comparar sus fortalezas y debilidades en términos de seguridad, economía, calidad de vida, cultura, accesibilidad a servicios públicos y oportunidades de trabajo y educación.
2. Realizar Data Pipeline, para poder implementar una base de datos, para montarlo en un servicio de cloud (google cloud).
3. Creación de Dashboards, KPIs y métricas que enseñen los principales resultados sobre los factores que influyen en las migraciones.
4. Desarrollar un modelo de machine learning que permita a los usuarios identificar sus prioridades y recibir recomendaciones personalizadas sobre el destino de emigración más adecuado para ellos.



5. Producir un repositorio en la plataforma Github, disponiendo el proyecto a la comunidad, para que esta expongan mejoras para futuras versiones del proyecto.

KPIs

- Cantidad de personas que entran a los países que acogen más argentinos en un período determinado es el "Número de migrantes". Este KPI simplemente cuenta el número total de personas que entran en un país en un período de tiempo específico, como un mes o un año. Otro KPI relacionado que se podría considerar es la "Tasa de migración", que mide la cantidad de migrantes en relación con la población total del país. Esto podría ser expresado como un porcentaje o como un número por cada 1,000 o 10,000 habitantes.
- Medir la cantidad de personas que abandonan Argentina en un período determinado es el "Número de emigrantes". Este KPI simplemente cuenta el número total de personas que abandonan un país en un período de tiempo específico, como un mes o un año. Otro KPI relacionado que se podría considerar es la "Tasa de emigración", que mide la cantidad de emigrantes en relación a la población total del país. Esto podría ser expresado como un porcentaje o como un número por cada 1,000 o 10,000 habitantes.
- Medir el impacto de la migración en el desarrollo económico de los países de origen y destino incluyen:
 1. Participación laboral de los migrantes: mide el porcentaje de migrantes que están activos en el mercado laboral y contribuyen a la economía.
 2. Remesas: mide el dinero que los migrantes envían a sus países de origen, que puede ser una importante fuente de ingresos para las economías locales.



3. Crecimiento económico: mide el cambio en el PIB (Producto Interno Bruto) de los países de origen y destino, lo cual puede ser afectado por la migración.
4. Desempleo: mide el porcentaje de la población que está buscando trabajo y no puede encontrarlo, lo cual puede ser afectado por la migración debido a la llegada o partida de trabajadores.

Solución propuesta

Stack tecnológico

Programación: Python, Jupyter notebook, Pandas, Scikit learn y Matplotlib.

Python cuenta con una amplia variedad de librerías de análisis de datos, como NumPy, Pandas y Matplotlib, que nos permiten manipular y visualizar datos de manera fácil y eficiente. También, Scikit-learn incluye una amplia variedad de algoritmos de aprendizaje automático.

Desarrollo Web: HTML, CSS, JavaScript y Flask API.

HTML, CSS y JavaScript son los estándares de la industria para el desarrollo web. Flask API es rápido, eficiente y fácil de utilizar, además, se integra fácilmente con python.

Cloud Computing: Google Cloud.

Google Cloud se integra fácilmente con otras tecnologías que utilizaremos, como Google Storage, Google BigQuery y Google Composer.

Flujo de Trabajo: Apache Airflow.

Nos permite automatizar flujos de trabajo complejos, lo que significa que podemos ejecutar tareas repetitivas de manera rápida y eficiente.

Contenerización: Docker.

Docker es una de las tecnologías de contenedores más populares y ampliamente utilizadas, y Kubernetes está diseñado para trabajar con Docker de manera nativa, lo que significa que la integración entre ambas tecnologías es fluida y sin problemas.



Database: MySQL.

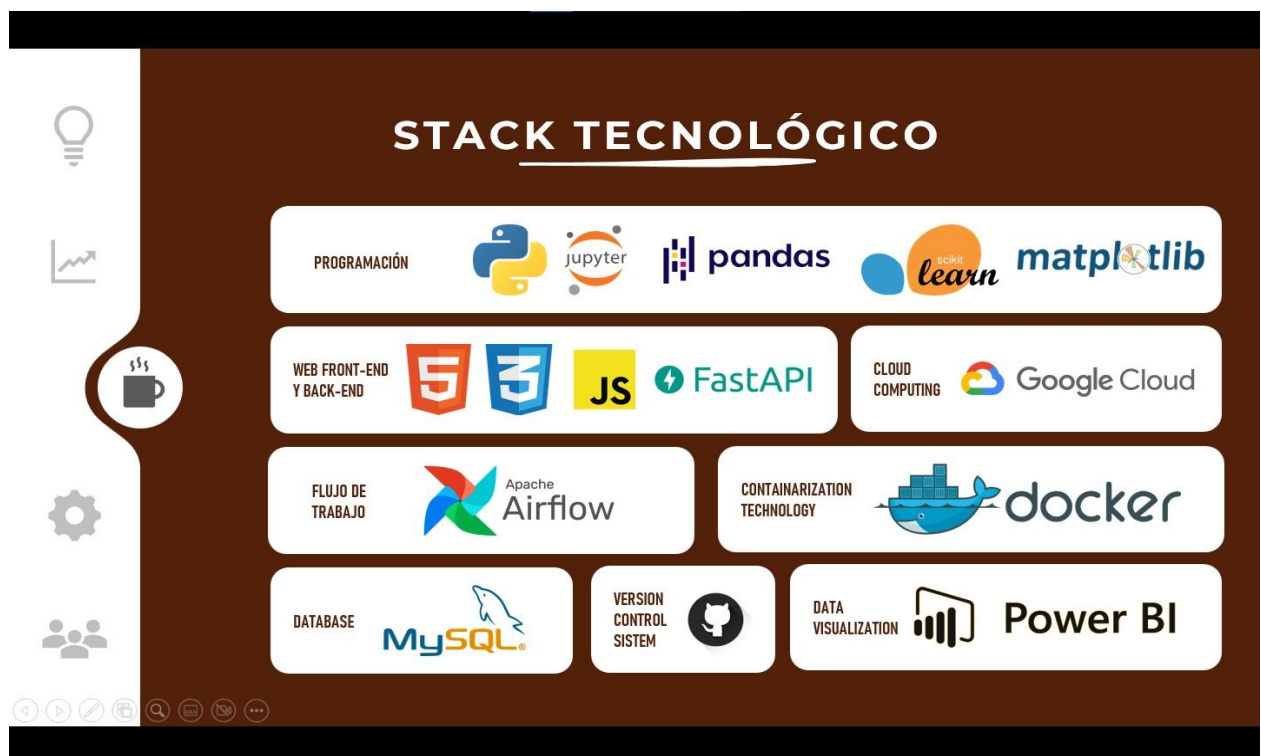
Principalmente por su escalabilidad, por lo fácil que se integra a con otras tecnologías y por la cantidad y calidad de su documentación.

Herramienta de visualización: Power BI.

Posee una amplia gama de herramientas de análisis y visualización avanzadas y además es fácil de utilizar.

Administrador de versiones: GitHub.

Nos permite colaborar en equipo, compartir nuestros códigos y mantener un registro de los cambios realizados en el código y revertir a versiones previas si es necesario.



Metodología de trabajo

- Se usará la metodología SCRUM la cual consiste en llevar a cabo un conjunto de tareas de forma regular con el objetivo de trabajar en forma conjunta.
- Progreso del proyecto dividido en semanas.
- Encuentros diarios del equipo: 8:30 hs ARG por discord
- Encuentros diarios con el Henry Mentor (Marcos Audisio) 13:00 hs ARG por Meet



- Encuentros semanales con el Product owner (Julio)
- Comunicación oficial con HM por Slack
- Comunicación con el equipo por Whatsapp
- Avances del proyecto en repositorio de github

Entregables

- **Semana 1.** Propuesta de trabajo
- **Semana 2.** Documentación
 - Stack elegido y fundamentación
 - Flujo de trabajo
- **Semana 3.**
 - Dashboards
 - Modelos de Machine Learning

Equipo y roles de trabajo

Nombre	Rol de trabajo
Juan Garcia	Data Engineering
Gian Rivas	Data engineering / DevOps
Camilo Ardila	Machine Learning Engineer
Matias Garro	ML engineer / Data Analyst
Camilo Pedreros	Data Analyst



Cronograma general - Gantt

Diagrama de Gantt para proyecto de flujos migratorios en Argentina																																	
1																																	
2																																	
3			Inicio del proyecto		Lunes, 6/02/2023	Lun, 06/02/2023				Lun 13/02/2023				Lun, 20/02/2023				Lun, 27/02/2023															
4			Fin del proyecto		Viernes, 3/03/2023	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	1	2	3	4	5
5	Tarea	Responsable	Progreso	Inicio	Fin	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D	L	M	M	J	V	S	D
6	Semana 1		100%	6/2/2023	12/2/2023																												
7	Busqueda datasets complementarios	Juan Garcia	100%	7/2/2023	9/2/2023																												
8	Análisis preeliminar de datasets	Gian Rivas	100%	9/2/2023	10/2/2023																												
9	Creacion documento propuesta del proyecto	Camilo Ardila	100%	8/2/2023	9/2/2023																												
10	Diseño de la estructura del proyecto y stack tecnologico	Matias Garro	100%	8/2/2023	9/2/2023																												
11	Creacion diagrama Grantt- colaboracion en el entregable	Camilo Pedreros	100%	8/2/2023	10/2/2023																												
12	Semana 2		0%	13/2/2023	19/2/2023																												
13	ETL/ diseño del modelo relacional	Juan Garcia	0%	13/2/2023	15/2/2023																												
14	Pipelines para el DW/Automatizacion Airflow	Gian Rivas	0%	13/2/2023	14/2/2023																												
15	Documentacion/ apoyo DW	Camilo Ardila	0%	14/2/2023	17/2/2023																												
16	DW/ Validacion de datos	Matias Garro	0%	15/2/2023	16/2/2023																												
17	diccionario datos/ apoyo documentacion	Camilo Pedreros	0%	16/2/2023	17/2/2023																												
18	Semana 3		0%	20/2/2023	26/2/2023																												
19	Documentacion/ Apoyo modelos ML	Juan Garcia	0%	21/2/2023	23/2/2023																												
20	Documentacion/ apoyo Dashboard (conexión a DB)	Gian Rivas	0%	22/2/2023	24/2/2023																												
21	Modelos ML en producción	Camilo Ardila	0%	23/2/2023	24/2/2023																												
22	Implementacion pagina web/ Apoyo dashboard	Matias Garro	0%	24/2/2023	24/2/2023																												
23	Diseño y elaboracion Dashboard y KPIs/ informe analisis	Camilo Pedreros	0%	24/2/2023	24/2/2023																												
24	Semana 4		0%	27/2/2023	5/3/2023																												
25	Readme general	Juan Garcia	0%	28/2/2023	1/3/2023																												
26	Ajustes finales	Gian Rivas	0%	28/2/2023	1/3/2023																												
27	Documentacion final	Camilo Ardila	0%	27/2/2023	1/3/2023																												
28	Presentacion	Matias Garro	0%	27/2/2023	28/2/2023																												
29	Storytelling	Camilo Pedreros	0%	28/2/2023	28/2/2023																												

Análisis Preliminar de datos

Los datos extraídos son de la página del Banco Mundial, la cual es una fuente confiable. Los datasets presentan datos históricos desde 1950 hasta 2021, los cuales presentan varios nulos y datos sobrantes para el estudio, principalmente en los años de 1950 a 1990. La idea es trabajar con la mayoría de los datos que no sean null por lo cual se quiere hacer el estudio en el rango de tiempo de 1990 hasta 2021, que es el rango de tiempo en donde se tienen datos completos.