



TRABAJO DE GRADO
Opción Seminario-Diplomado.

Análisis de datos sobre la migración colombiana de 25 a 40 años

Corporación Universitaria Remington.
Facultad de Ingeniería
Ingeniería de Sistemas

Elisabet Castaño Salazar
John Fredy Mira Mejía
Opción de Trabajo de grado Seminario-Diplomado.
2024.

Tabla de Contenidos

Resumen.....	4
Marco conceptual y contextual	4
Marco contextual	5
Marco conceptual.....	5
1. Concepto de migración	5
2. Análisis de datos	6
3. Visualización de datos	6
Desarrollo e implementación del aprendizaje.....	7
Datos	7
4. ¿Qué es un dato?	7
5. Tipos de datos	7
5.1. Cuantitativos	7
5.2. Cualitativos	7
6. Categoría de datos.....	7
6.1. Datos estructurados	7
6.2. Datos semiestructurados	7
6.3. Datos no estructurados	7
Visualización de datos	8
7. ¿Qué es un gráfico?.....	8
8. Objetivo.....	8
9. Tipos de gráficos.....	8
9.1. Columnas y barras.....	8
9.1.1. Barras	8
9.1.2. Columnas	9
9.1.3. Columnas o barras agrupadas	9
9.1.4. Columnas o barras agrupadas al 100%	10
9.2. Circular	10
9.2.1. Pastel	11
9.2.2. Anillo	11
9.3. Gráfico de líneas	12
10. Tablas dinámicas.....	13
11. Herramientas de visualización	13
12. Análisis de datos	15
12.1. Tipos de análisis.....	15
12.1.1. Análisis descriptivo.....	15
12.2. Analizando nuestro conjunto de datos	15
Conclusiones	16
Referencias.....	18

Resumen

Este proyecto se centra en analizar y visualizar los datos relacionados con los perfiles de los colombianos en edades de 25 a 40 años que han decidido migrar a otros países, usando técnicas de análisis y visualización de datos, para entender y sacar conclusiones de la tabla de registros extraída de los registros públicos de Colombia.

Se utilizan datos generales de colombianos en el exterior provenientes de la plataforma nacional de datos abiertos de Colombia. Se analiza las tendencias migratorias recientes, así como la fuga de talentos.

Se examinan aspectos educativos, civiles y étnicos de esta población migrante a través del análisis de datos y emplean técnicas de visualización para presentar los hallazgos de la manera más clara y comprensible posible.

Palabras clave

Migración, Análisis de datos, Visualización de datos, Población joven, Colombia

Marco conceptual y contextual

Marco contextual

El enfoque del presente informe pretende analizar las características, tendencias y patrones migratorios sobre colombianos registrados en el extranjero, específicamente aquellos que se encuentran en el rango de edad de 25 y 40 años. El crecimiento notable en el índice de migración desde Colombia hacia otras naciones es un fenómeno que refleja las aspiraciones individuales de los ciudadanos en la búsqueda de oportunidades en el extranjero como el contexto socioeconómico del país.

El informe se basará en datos estadísticos relevantes y confiables, obtenidos de la plataforma nacional de datos abiertos de Colombia, la cual contiene datos públicos que se pueden usar para analizarlos, ya que son obtenidos de una fuente oficial y conocida, que proporciona información actualizada sobre diversos aspectos y sectores del país.

Se llevará a cabo un análisis de los datos recientes sobre la migración colombiana, incluyendo destinos preferenciales, características demográficas de la población migrante y el impacto que este fenómeno puede tener frente aspectos económicos, sociales y culturales. Se explorarán temas como la fuga de talentos, la dinámica de la población en edad productiva, así como los desafíos y oportunidades que este fenómeno plantea para el desarrollo social y económico de Colombia.

En este contexto, resulta importante comprender las motivaciones y características de los colombianos que han optado por buscar nuevas oportunidades en el extranjero, para poder llegar a conclusiones sobre los perfiles de las persona que toman la decisión de emigrar, y enfocar soluciones en esta población. A través del análisis de datos generales, este informe se propone proporcionar una visión detallada del perfil educativo, civil y étnico de esta población migrante.

Marco conceptual

1. Concepto de migración

La migración es un fenómeno humano que implica el desplazamiento de individuos de un lugar a otro, ya sea dentro de las fronteras de un mismo país, o hacia otro internacional, con la intención de establecerse de forma permanente o temporal. Este movimiento puede estar motivado por una variedad de factores complejos que van desde aspectos económicos, sociales, políticos y ambientales.

- Factores que impulsan la migración

- **Económicos:** Los individuos pueden migrar en busca de mejores oportunidades laborales, acceso a mayores ingresos, mejora general en su calidad de vida.

- **Sociales:** La migración también puede estar impulsada por la necesidad de la huida de situaciones de violencia o persecución, así como la búsqueda de mejores condiciones de vida para si mismo y los seres queridos.
- **Políticos:** Los conflictos armados, regímenes autoritarios, la persecución por motivos religiosos o ideológicos pueden forzar a las personas a abandonar sus lugares de origen en busca de seguridad y libertad.
- **Ambientales:** Factores como desastres naturales, el cambio climático y la degradación ambiental pueden desencadenar movimientos migratorios debido a la pérdida de recursos, la amenaza a la seguridad y la viabilidad de la vida en determinadas áreas.

2. Análisis de datos

El análisis de datos es el proceso de recopilación, limpieza, transformación y análisis de conjuntos de datos con el objetivo de identificar tendencias, patrones y relaciones significativas. Este proceso permite extraer información valiosa para la toma de decisiones y la aplicación de conocimientos en una amplia gama de temas.

En este proyecto, el análisis de datos se centrará en los colombianos registrados en el extranjero, específicamente aquellos dentro del rango de edad de 25 a 40 años. Se buscará comprender las dinámicas migratorias, identificar factores determinantes y proponer soluciones basadas en evidencia para abordar los desafíos asociados con este fenómeno.

3. Visualización de datos

La visualización de datos es una técnica que consiste en representar gráficamente información de manera clara, concisa y efectiva. A través de ésta es posible comunicar patrones, relaciones y tendencias mediante estímulos visuales, fomentando la comprensión y el análisis de los datos obtenidos.

Con el objetivo de presentar los resultados del análisis de la manera más clara y accesible para la mayor cantidad de personas, en el presente trabajo, se utilizarán diversas técnicas de visualización de datos. Lo anterior se materializará en diversos gráficos que diseñarán una imagen clara de los descubrimientos del análisis, permitiendo la interpretación y comunicación de los resultados de forma sencilla y asequible.

Desarrollo e implementación del aprendizaje

Datos

4. ¿Qué es un dato?

Es una representación de un atributo o variable, que nos permite analizar un evento o hecho. También se puede definir como una unidad de información que puede ser procesada y analizada.

5. Tipos de datos

5.1. Cuantitativos

Un dato cuantitativo es aquel que se puede medir o contar, como un número (QuestionPro.com). Por ejemplo, en el caso de análisis del presente proyecto, se tienen datos cuantitativos como son la edad de las personas y la estatura.

5.2. Cualitativos

Un dato cualitativo es el que se puede categorizar por una cualidad (QuestionPro). Se expresa en palabras y no se puede medir ni contar. En el caso del presente proyecto se tiene el género, el estado civil, la etnia, el área de conocimiento, la subárea de conocimiento, el nivel académico y el país al cual migro la persona.

6. Categoría de datos

6.1. Datos estructurados

Son aquellos que ya están organizados y con un formato, ya sea numérico, de fecha, de texto. Estos datos siguen un esquema ya definido y que ayuda a procesarlos fácilmente. Son más fáciles de filtrar, ayudando a realizar búsquedas y análisis rápida y eficientemente. Este tipo de datos suelen estar almacenados en bases de datos y hojas de cálculo.

6.2. Datos semiestructurados

Estos datos se encuentran medianamente organizados, pero no siguen un esquema definido como en el caso de los estructurados. Suelen ser datos que llevan un mensaje o un comentario, y en los cuales podemos analizar los sentimientos de estos.

6.3. Datos no estructurados

Son datos que no se encuentran organizados, ni tienen un formato definido, tampoco siguen un esquema. Son difíciles de procesar por la falta de organización y formato, por lo cual requiere diferentes técnicas de procesamiento como analizar el lenguaje o el texto. Para realizar búsquedas y análisis es más complejo por esta misma razón, suelen ser datos que están en imágenes, videos, audios, documentos de texto como Word, PDF, es importante convertir estos datos a estructurados, para poder analizarlos.

Visualización de datos

7. ¿Qué es un gráfico?

Es una representación de datos visual que ayuda a analizar de una manera más sencilla relaciones, tendencias y patrones entre los datos. Da un apoyo para poder contar la historia detrás de estos, a demostrar las conclusiones o decisiones que se quieren o se van a tomar.

Con estos gráficos se pueden realizar paneles que ayudan analizar los datos de maneras diferentes, ya que normalmente estos paneles tienen diversas representaciones visuales que permiten comunicar la información.

8. Objetivo

El objetivo principal de estos gráficos es mostrar la información de una manera que se pueda entender mejor que lo que se puede ver a primera vista en las bases o tablas de datos donde almacenamos estos. Para cumplir este objetivo se debe tener en cuenta las siguientes características al momento de crear un gráfico:

- **Auto explicativo:** Deben ser simples, estos se usan tanto en presentaciones como en informes, por lo que quien vea el informe no debe requerir la presencia del autor del gráfico, debe por sí mismo entender el contenido y sacar conclusiones propias de este.
- **Indicar las unidades de valores:** Para poder entender en su totalidad los gráficos, se debe dar la información suficiente para hacerlo, para esto se deben aclarar los datos que estamos mostrando, especificar sus unidades, si representa años, meses, miles o millones.
- **Nombre de la leyenda:** Igualmente que debe aclarar las unidades de los datos, se debe aclarar a que información pertenecen estas unidades, si es a edades, años, litros, categorías.
- **Presentación limpia:** Al ser información que se quiere representar visualmente, se deben encontrar los colores, tamaños y fuentes, ideales y amigables para hacerlo, que se pueda dar enfoque a lo que queremos sin exagerar ni lastimar la visión. Se deben retirar los elementos de distracción que puedan quitar la importancia a los datos.

9. Tipos de gráficos

9.1. Columnas y barras

Las gráficas de barras o columnas sirven para comparar diferentes categorías de datos, visualizando las tendencias o cambios a lo largo del tiempo (TuDashboard) o rango de unidad.

9.1.1. Barras

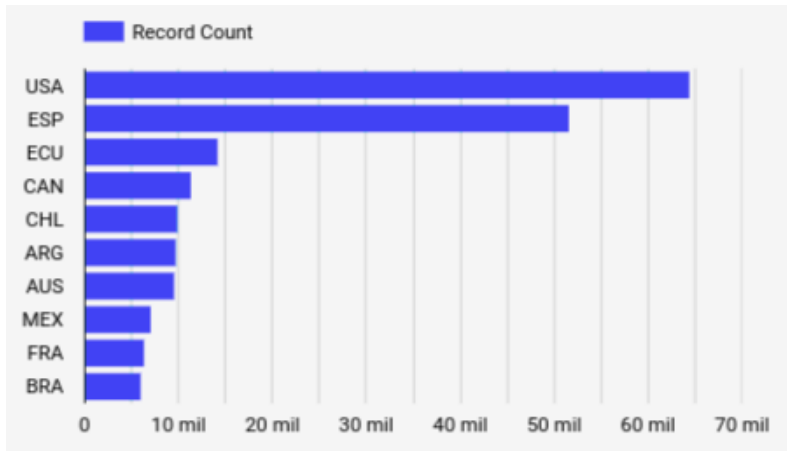


Ilustración 1. Gráfico de barras, países con colombianos registrados.

En la Ilustración 1 se pueden ver agrupados por conteos los colombianos que han decidido migrar a otro país, logrando visualizar que el país en el cual hay más colombianos es en Estados Unidos, y el segundo con mayor número es España, el resto de los países no es una opción tan popular como estos dos primeros.

9.1.2. Columnas

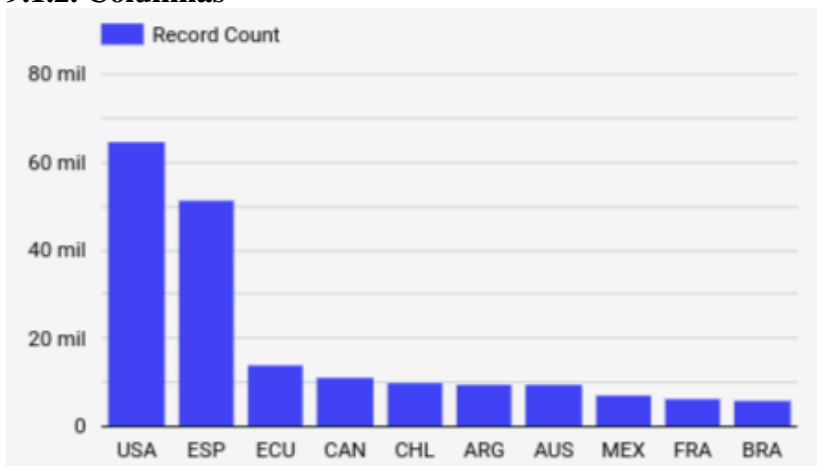


Ilustración 2. Gráfico de columnas, países con colombianos registrados.

Los mismos gráficos de barras pueden visualizarse por columnas, dependiendo del número de barras o columnas que se desplegarán, podemos elegir entre sí, cuando son demasiadas columnas o barras, se elige la opción del gráfico de barras, ya que permite expandir la gráfica verticalmente para visualizar mejor la información.

9.1.3. Columnas o barras agrupadas

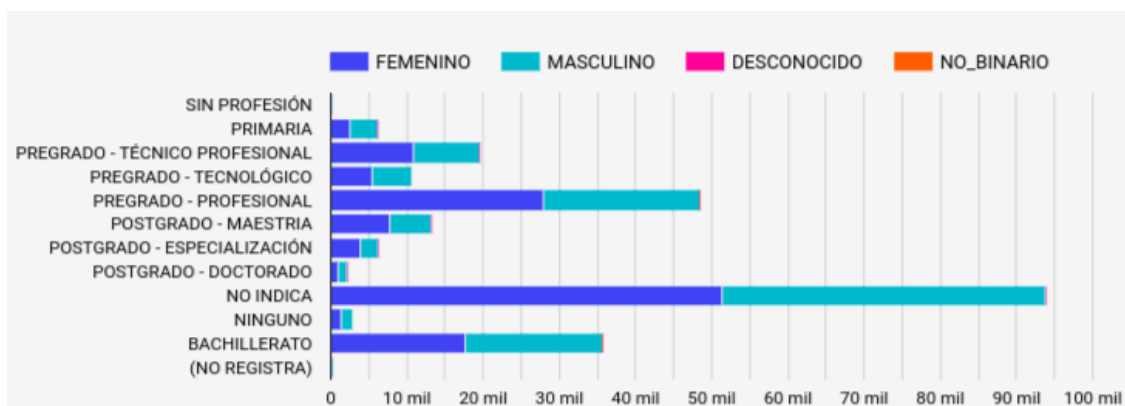


Ilustración 3. Gráfico de columnas agrupadas por género y nivel académico.

Un tipo de gráfica tanto en columnas como en barras, son las agrupadas, que ayudan a visualizar un conjunto de datos teniendo en cuenta dos valores. En la Ilustración 3 se agrupan a los colombianos por género y nivel académico, el cual muestra por cada barra como se divide entre los diferentes géneros, por ejemplo, los colombianos se dividen mayormente entre el género masculino y femenino, ya que podemos visualizar las barras de color azul oscuro y claro ocupando una gran parte de las barras, también como la barra de no indica nivel académico pertenece al de mayor número de colombianos y como se divide en gran cantidad por el género femenino y el segundo número el masculino.

9.1.4. Columnas o barras agrupadas al 100%

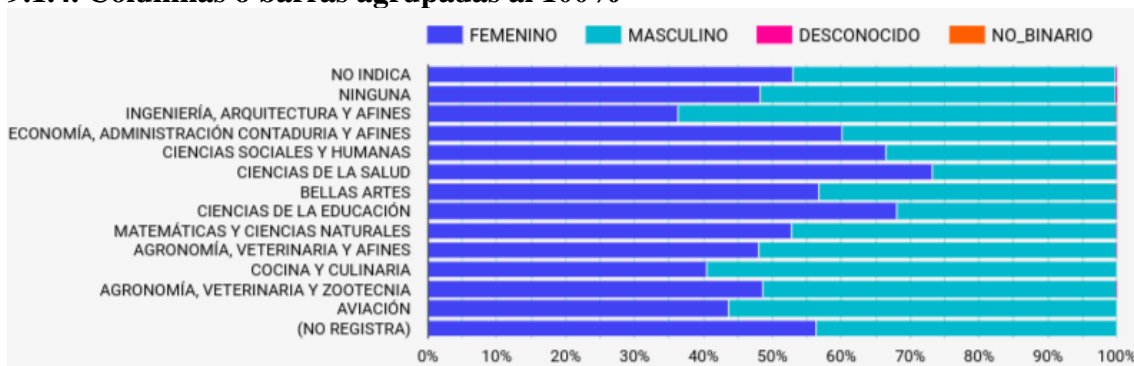


Ilustración 4. Gráfico de columnas agrupadas al 100% por área de conocimiento y género.

Parecida a la gráfica de columnas o barras agrupadas, tenemos la gráfica de columnas o barras agrupadas al 100%, donde se representa los datos por medio de porcentajes, se puede visualizar en la Ilustración 4, como se divide por porcentajes el género en cada barra de área de conocimiento, como los géneros con mayor porcentaje son el masculino y femenino, y como entre estos dos géneros los porcentajes no llegan a ser menores del 30% o mayor al 80%

9.2. Circular

Las gráficas circulares representan el tamaño de los datos por porciones respecto al total de estos. (TuDashboard)

9.2.1. Pastel

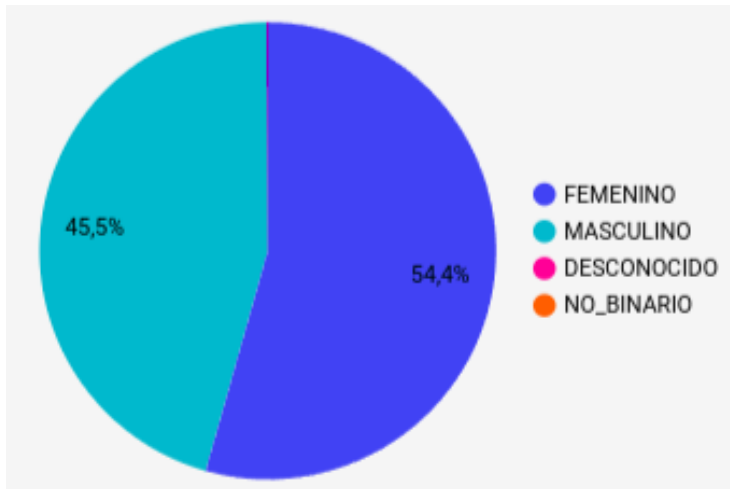


Ilustración 5. Gráfico circular de pastel, dividido por género.

En la Ilustración 5, se puede visualizar un círculo completo, el cual se divide por porciones respecto al porcentaje total, por ejemplo, el porcentaje que ocupa el género femenino, masculino, no binario, y de las personas de las cuales no se conoce su género, con respecto al total de los datos, logrando visualizar que el mayor porcentaje de personas que deciden residir en otro país es el género femenino.

9.2.2. Anillo

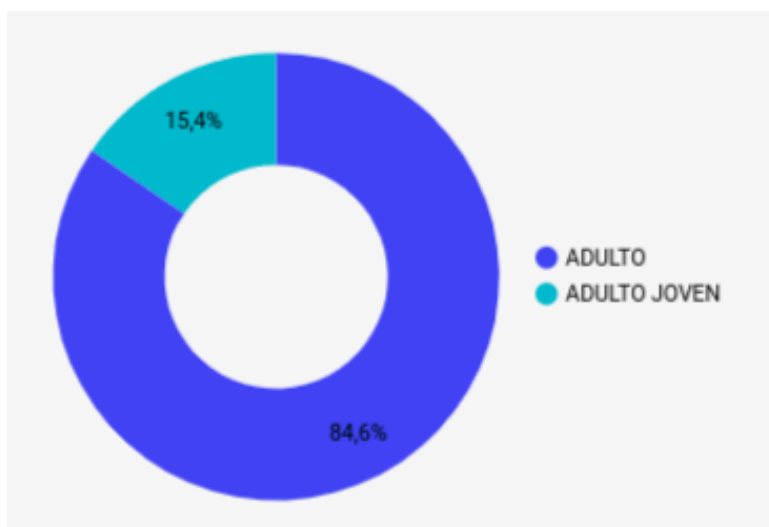


Ilustración 6. Gráfico circular de anillo, dividido por grupo de edad.

La gráfica de anillo, funciona igualmente que la de pastel, mostrando por porcentajes el los valores en conjunto del total de los registros, en la Ilustración 6, se visualiza por ejemplo, el porcentaje que ocupan los adultos jóvenes y los adultos, logrando ocupar el mayor porcentaje los adultos.

9.3. Gráfico de líneas

Este tipo de gráfico ayuda a identificar tendencias, mayormente utilizado para identificar las evoluciones de nuestros datos de estudio referentes al tiempo (IBM Docs, Cognos Analytics), ayudando a visualizar picos altos o bajos, ya sea de ventas, inversiones, etc.

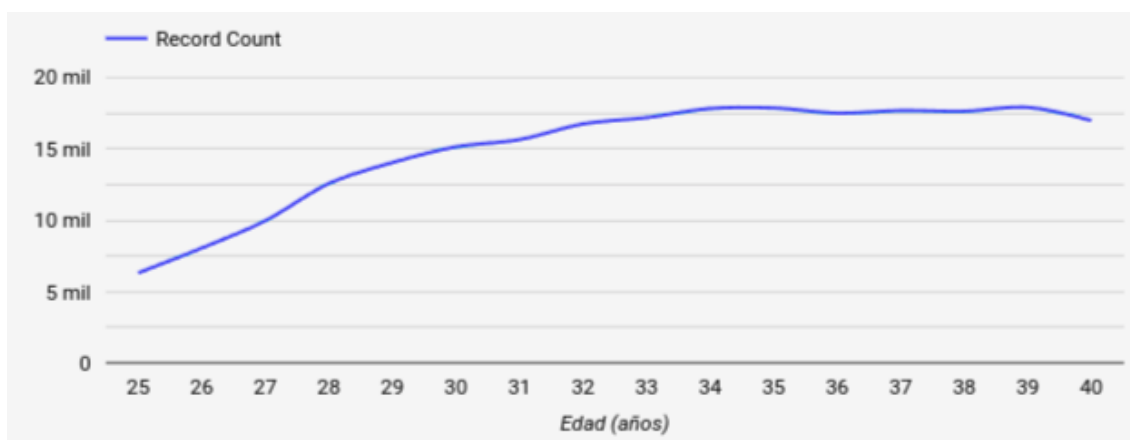


Ilustración 7. Gráfico de líneas, tendencia de las edades de los colombianos en otros países.

En la Ilustración 7 se logra visualizar como la cantidad de los datos varía entre las diferentes edades que encontramos en estos, logrando ver que las edades entre 25 y 30 ocupan la menor cantidad de los datos, y como las otras edades ocupan un mayor número de estos.

10. Tablas dinámicas

Las tablas dinámicas son una herramienta para calcular, resumir y analizar datos (Microsoft support) que se obtienen de otro conjunto de datos de mayor tamaño.

No hay reglas para usarlas, como una estructura, ya que se pueden organizar de la manera que sea más útil para representar o filtrar los datos que se requieren para el análisis de datos.

Permiten realizar operaciones como conteos, promedios, sumas sin necesidad de usar formulas, y a cambiar filas, columnas y filtros fácilmente.

Etnia de la persona	Código ISO país / Record Count						Total
	USA	ESP	ECU	CAN	CHL	ARG	
NINGUNA	28.997	26.195	5.889	6.348	4.951	5.216	118.978
OTRO	31.466	20.819	6.033	4.384	3.498	3.352	101.899
SIN ETNIA REGISTRADA	2.734	2.424	1.381	446	365	843	10.209
AFRODESCENDIENTE	1.024	1.795	744	142	996	235	6.397
INDÍGENA	216	280	170	11	77	60	1.152
RAIZAL DEL ARCHIPIELAGO DE SAN ANDR...	52	21	4	6	5	12	159
GITANO	31	29	15	1	6	1	126
PALENQUERO DE SAN BASILIO	18	7	4	-	2	3	48
Total	64.538	51.570	14.240	11.338	9.900	9.722	238.968

Ilustración 8. Tabla dinámica, visualización de etnia en los diferentes países migratorios.

Como se visualiza en la Ilustración 8, las filas se dividen en la etnia de la persona, y en las columnas se pueden ver los países a los cuales estas personas deciden migrar, cada celda que es la unión de la fila y la columna nos muestra el total de registros en esta categoría, por ejemplo, en el país de Estados Unidos los colombianos con ninguna etnia abarcan un total de 28.997 personas, también se puede ver en la última columna un total de las personas que no tienen ninguna etnia, y una fila final que nos contabiliza el total de personas en cada país.

11. Herramientas de visualización

Existen diversas aplicaciones o softwares que incluyen herramientas y funcionalidades para crear las gráficas y tablas que se mostraron en los puntos anteriores, para ello estas herramientas permiten ingresar un conjunto de datos, transformarlos y manipularlos visualmente.

En este proyecto se usó la herramienta Data Studio actualmente conocida como Looker Studio, la cual es una herramienta en línea gratuita de Google que ayuda a realizar informes y paneles fácilmente (Bloo.Media). Permitiendo subir los datos desde diferentes fuentes, como son archivos CSV, Excel, conexión a una base de datos, etc.

Con esta herramienta se realizaron dos paneles que ayudan a visualizar los datos generales de los colombianos en el extranjero e igualmente su nivel educativo.

La Ilustración 9 visualiza la información general de los colombianos en el exterior, cuantos se clasifican como adultos jóvenes estando en el rango de 25 a 28 años, cuantos están en la clasificación de adultos en edades comprendidas entre los 29 a 40 años, cuantos de estos colombianos son solteros. También se visualizan gráficas como las de barra mencionadas en el punto 3.1.1, otra de pastel mencionada en el punto 3.2.1, la de anillo mencionada en el punto 3.2.2 y finalmente la de línea mencionada en el punto 3.3.

En la Ilustración 9 se observa que el panel contiene solo información general, sin saturar el panel de información logrando que sea sencilla y entendible para quien vea el informe.

Información general colombianos en el exterior

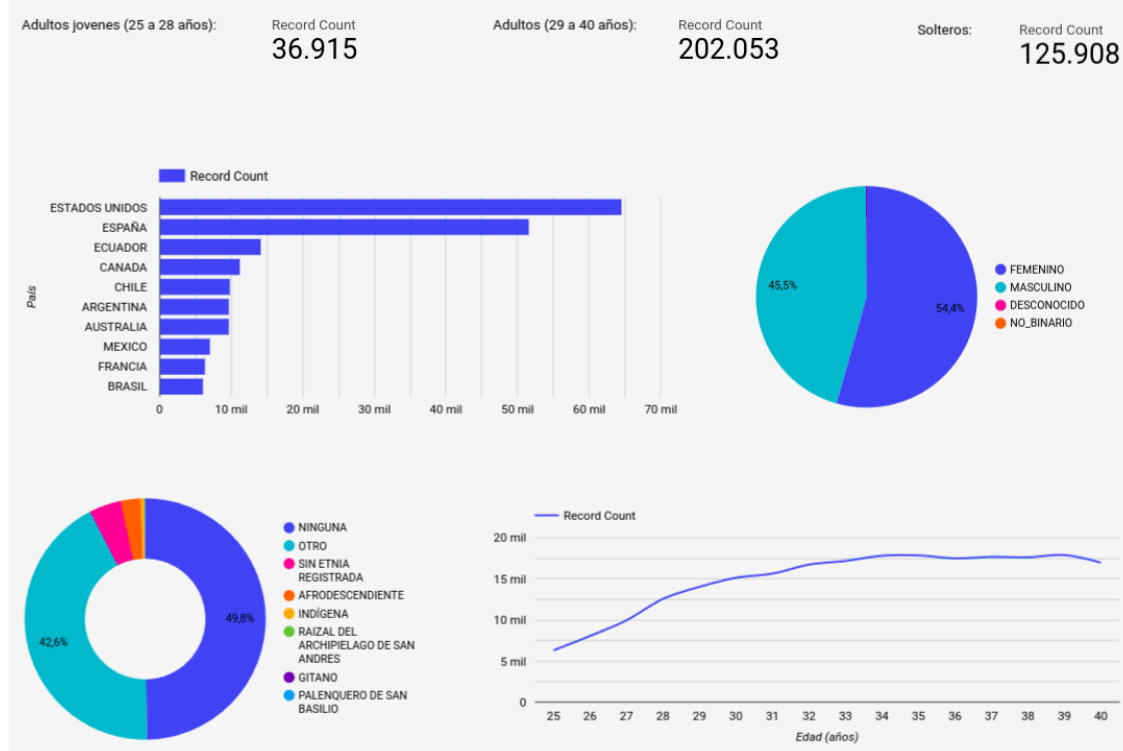


Ilustración 9. Panel de información general colombianos en el exterior

En la Ilustración 10 se visualiza el nivel educativo de los colombianos, comprendiendo esto en tres secciones con gráficas, un gráfico de barras con el nivel educativo, mostrando cuantos colombianos llegaron a cierto nivel educativo, como se observa muchos optan por no especificar este, y también que una gran cantidad de ellos tiene un pregrado profesional, se visualiza otro gráfico de pastel que muestra las subáreas de conocimiento, y una tabla de calor, donde los colores fuertes nos indican donde están concentrados los números

mayores, que como se ilustra en Estados Unidos se ubica un gran número de colombianos con área de conocimiento en ingeniería, arquitectura y afines.

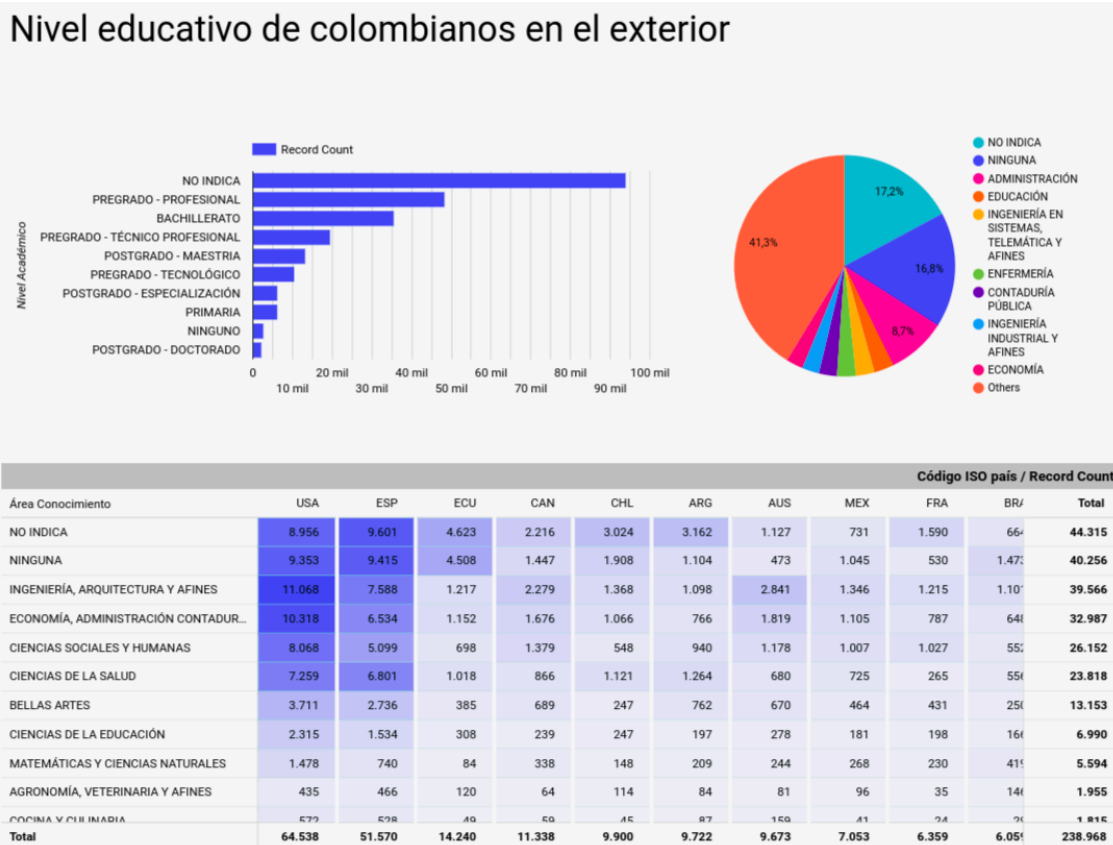


Ilustración 10. Panel de nivel educativo de colombianos en el exterior

12. Análisis de datos

Es el proceso en el cual se exploran, transforman y examinan los datos para identificar las tendencias y patrones (Alteryx.com) que podemos visualizar en los gráficos y paneles para con estos respaldar decisiones, descubrir información útil y comunicar hallazgos.

12.1. Tipos de análisis

12.1.1. Análisis descriptivo

El análisis descriptivo sirve para describir un conjunto de datos considerando lo que ha sucedido en el pasado.(QuestionPro, Marytere Narvaez) Permite describir las características del conjunto de datos.

12.2. Analizando nuestro conjunto de datos

Con los paneles de la Ilustración 9 y la Ilustración 10 se puede llegar a varias conclusiones sobre los datos que se recolectaron de los colombianos que residen en el exterior:

- El país preferente para el cual los colombianos deciden migrar es a Estados Unidos, siendo esta una potencia mundial, que ofrece oportunidades laborales y económicas a niveles que Colombia actualmente no tiene, haciendo la calidad de vida mucho mejor que la que se puede llegar a tener en nuestro país.
- El índice de adultos que deciden migrar a otros países es mucho mayor que el de adultos jóvenes, esto podemos predecir que es porque los adultos llevan el plan de migrar a otro país con mucho más tiempo, teniendo en cuenta los factores económicos, papeleo, etc. que tienen que realizar para poder tener una permanencia allí.
- Se puede visualizar que un poco más de la mitad de estos colombianos son solteros, considerando que es sencillo decidir migrar cuando no se tienen ataduras al país, como tener pareja o estar casado, tener hijos.
- Por el lado educativo, se puede visualizar como en forma general se pierde un gran porcentaje de profesionales y talentos en diferentes áreas de conocimiento, lo cual muestra la falta de oportunidades y remuneración en el país para dichos profesionales, llevándolos a la decisión de buscar estas oportunidades en un país extranjero, en el cual las oportunidades son mayores, logrando avanzar en su vida profesional, y la remuneración es mejor, logrando un aporte económico mejor para aumentar su calidad de vida y apoyar la de su familia en Colombia.

Conclusiones

Se puede concluir que el crecimiento de la población migrante va a continuar, teniendo en cuenta que hay edades en la cual este fenómeno se presenta con mayor frecuencia, mostrando que algunas de las razones más comunes de tomar esta decisión es la falta de oportunidades y la falta de calidad de vida que se puede aspirar en Colombia.

Este proyecto se enfocó en visualizar por medio de gráficas los perfiles de los colombianos que han migrado, ayudando a enfocarnos en estos perfiles para evitar este fenómeno que, aunque no es del todo malo, es importante lograr disminuirlo, para que el talento aporte al crecimiento del país y no se vea en la necesidad de buscar oportunidades en otros lugares.

Analizar este tipo de situaciones apoyados de gráficos, muestra cómo se puede llegar a conclusiones sin la necesidad de tener la lista de datos completos, si no que con un poco de contexto y organización entender lo que se quiere comunicar, mostrando la gran importancia que tiene la visualización de datos en la actualidad, teniendo en cuenta la magnitud de datos que están almacenados y que cada día está en constante crecimiento.

Referencias

Colombianos en el exterior de 25 a 40 años. Sitio web:

<https://www.datos.gov.co/Estadisticas-Nacionales/Colombianos-en-el-exterior-de-25-a-40-a-os/3k8j-4d42>

Alteryx, Análisis de datos. Sitio web:

<https://www.alteryx.com/es/glossary/data-analytics#:~:text=El%20an%C3%A1lisis%20de%20datos%20es,respaldar%20la%20tomada%20de%20decisiones>

QuestionPro, Análisis de datos. Sitio web:

<https://www.questionpro.com/es/analisis-de-datos.html>

IMB, documentación, tipos de gráficos. Sitio web:

<https://www.ibm.com/docs/es/cognos-analytics/11.1.0?topic=charts-chart-types>

TuDashboard, tipos de gráficas. Sitio web:

<https://tudashboard.com/tipos-de-graficas/>

Microsoft support, crear una tabla dinámica para analizar datos. Sitio web:

<https://support.microsoft.com/es-es/office/crear-una-tabla-din%C3%A1mica-para-analizar-datos-de-una-hoja-de-c%C3%A1lculo-a9a84538-bfe9-40a9-a8e9-f99134456576>

Bloo medio, tutorial Google Data Studio. Sitio web:

<https://bloo.media/blog/tutorial-google-data-studio/#:~:text=Google%20Data%20Studio%20es%20una,Google%20y%20otras%20fuentes%20externas>

QuestionPro, tipos de análisis de datos. Sitio web:

<https://www.questionpro.com/blog/es/tipos-de-analisis-de-datos-2/>