

ANALISIS EXPLORATORIO DE DATOS (EDA):

Evolución de la demografía vasca y el impacto en el uso y conocimiento del Euskara.



INDICE

1. Introducción
2. Hipotesis
3. Descripción de los Datos
4. Metodología
5. Análisis de los Datos
6. Conclusiones
7. Bibliografía

1. Introducción

En este análisis exploratorio de datos (EDA) se aborda la evolución del conocimiento y uso del euskera en Euskadi durante el periodo 2001-2021. A través de los datos proporcionados, se busca identificar las tendencias de crecimiento en la población que habla euskera, la distribución territorial de los hablantes, y el impacto de la inmigración en la preservación del idioma. Además, se analiza el uso del euskera en diferentes contextos (hogar, trabajo, etc.) para entender su normalización en la sociedad vasca.

El idioma euskera es hablado en las comunidades autónomas de Euskadi y Navarra. Este análisis tiene como objetivo explorar la evolución de este idioma en el contexto de los cambios demográficos y sociales que enfrenta el País Vasco en los últimos años, caracterizados por una baja natalidad que se ve compensada, en parte, por la llegada de inmigrantes.

Algunas de las preguntas clave que guían este estudio son:

1. ¿Hay más hablantes de euskera en la actualidad?
2. ¿Cómo ha evolucionado la población en Euskadi desde 2001 hasta 2021, tanto en términos generales como por grupos de edad?
3. ¿Cuál es la relación entre el crecimiento o disminución de la población y el uso del euskera?
4. ¿Tiene la inmigración un impacto negativo en la preservación del euskera debido a la dificultad de aprendizaje del idioma?
5. ¿Cómo está afrontando el gobierno los cambios demográficos para proteger uno de los idiomas más antiguos de Europa?
6. ¿Existe interés en conservar y promover el uso del euskera?

Antes de profundizar en los datos que responderán estas preguntas, es importante señalar que, desde una perspectiva personal, el euskera ha enfrentado numerosos altibajos a lo largo de los años. Se trata de un idioma complejo, especialmente para quienes no lo aprenden desde la infancia, aunque no imposible de dominar. Existen casos de personas que, sin tener conocimientos previos, han logrado obtener el EGA, un título que acredita el dominio del euskera y permite acceder a diversos empleos en el ámbito público.

Es evidente que la situación del idioma varía significativamente entre las áreas urbanas y rurales. En las capitales, el castellano tiene un mayor peso, mientras que en las zonas rurales el euskera predomina. Además, hay diferencias notables entre las provincias; por ejemplo, Gipuzkoa tiene una mayor presencia del euskera en comparación con Álava.

En cuanto a la inmigración, aunque no se dispone de datos concretos sobre el porcentaje que representa dentro de la población vasca, es razonable pensar que los inmigrantes no suelen aprender el idioma. La adaptación a un nuevo país conlleva numerosos retos culturales y lingüísticos, y el euskera, siendo un idioma que no guarda similitudes con el castellano, no suele ser una prioridad.

El Gobierno Vasco, en su esfuerzo por impulsar el uso del euskera, implementó hace unos años el modelo educativo D, donde la enseñanza pública como privada o concertada se imparte exclusivamente en euskera. Aunque esta medida busca garantizar que las generaciones jóvenes hablen o comprendan el idioma, no parece ser la solución más efectiva. La realidad es que muchas familias no pueden educar a sus hijos exclusivamente en euskera, lo que ha llevado a la apertura de otras líneas educativas en castellano.

El euskera enfrenta retos constantes. A diferencia de otros idiomas regionales como el catalán, que guarda una mayor similitud con el castellano y es más accesible para los hablantes, el euskera es menos escuchado y, posiblemente, menos hablado hoy en día. Es fundamental reflexionar sobre las estrategias necesarias para garantizar su supervivencia y promover su uso en todos los ámbitos de la sociedad.

2. Hipotesis

• Hipótesis 1:

El uso del euskera ha aumentado proporcionalmente en Euskadi, impulsado por políticas lingüísticas y educativas

• Hipótesis 2:

La inmigración, especialmente de regiones donde el euskera no es hablado, ha afectado negativamente la adopción del idioma,

3. Descripción de los Datos

Los objetivos de este análisis son los siguientes:

- Analizar la evolución del conocimiento del euskera entre 2001 y 2021.
- Explorar las diferencias en la distribución del conocimiento del euskera por provincia (Gipuzkoa, Bizkaia, Álava).
- Identificar patrones relacionados con la inmigración y su impacto en la adopción del euskera.
- Estudiar el uso del euskera en diversos contextos sociales (hogar, trabajo, administración, etc.).
- Determinar posibles correlaciones entre los diferentes factores que afectan al conocimiento y uso del euskera.

Los datos utilizados provienen de Eustat (Instituto Vasco de Estadística), con registros desde 2001 hasta 2021 principalmente. Los datos están organizados por provincia (Gipuzkoa, Bizkaia, Álava). Los datasets incluyen las siguientes variables:

- **Población total:** Número total de habitantes por año y provincia.
- **Conocimiento del euskera:** Número de personas que saben euskera en cada provincia.
- **Uso del euskera en diversos ámbitos:** Datos sobre el uso del euskera en casa, con los hijos, en el trabajo, y en otros entornos institucionales.
- **Procedencia geográfica:** Desglosado por países de origen, con especial enfoque en la población autóctona y la inmigración.

4. Metodología

El análisis se realizó utilizando **Python** y las librerías **Pandas** para la manipulación de datos, **Matplotlib** y **Seaborn** para la visualización gráfica, y **Scipy** para el análisis de correlación. Los pasos principales del análisis fueron:

- **Limpieza de datos:** Se eliminaron valores nulos y se transformaron las variables categóricas a un formato adecuado para el análisis.
- **Análisis univariate:** Se analizaron las distribuciones de cada variable para identificar tendencias generales a lo largo del tiempo.
- **Análisis bivariate:** Se exploraron las relaciones entre el conocimiento del euskera y factores como la provincia, el año y la inmigración.
- **Correlación:** Se calculó la correlación entre el nivel de euskera y la inmigración para identificar patrones y dependencias.

5. Análisis de los Datos (EDA)

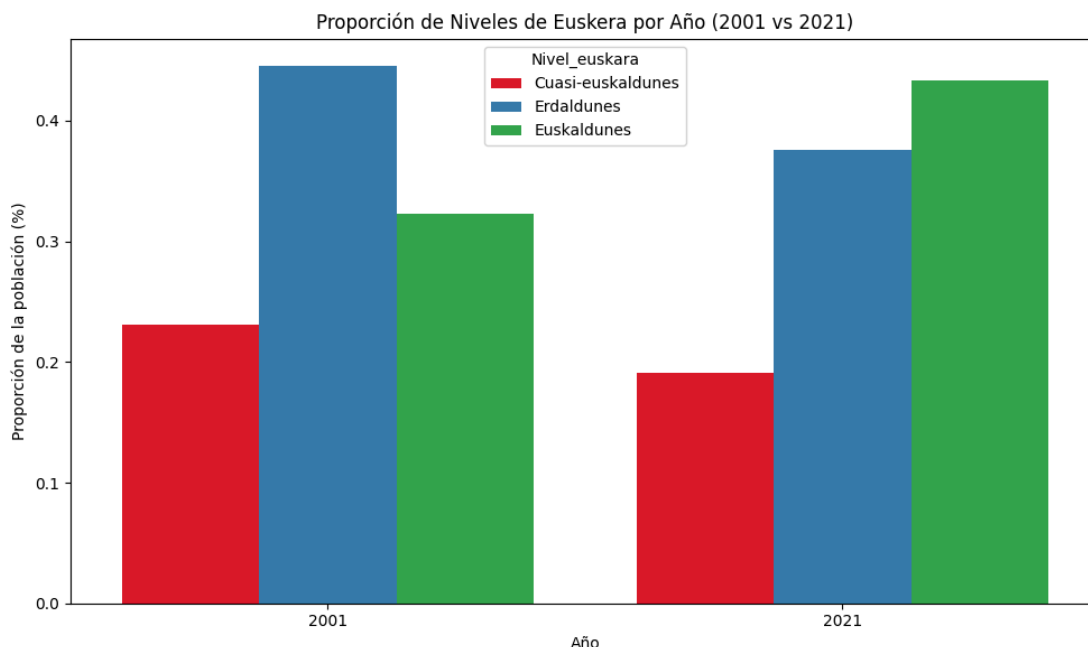
5.1 Análisis Nivel Euskera

Se analizaron las distribuciones de las variables relacionadas con el conocimiento del euskera.

Tenemos que saber como funcionan los niveles de euskera:

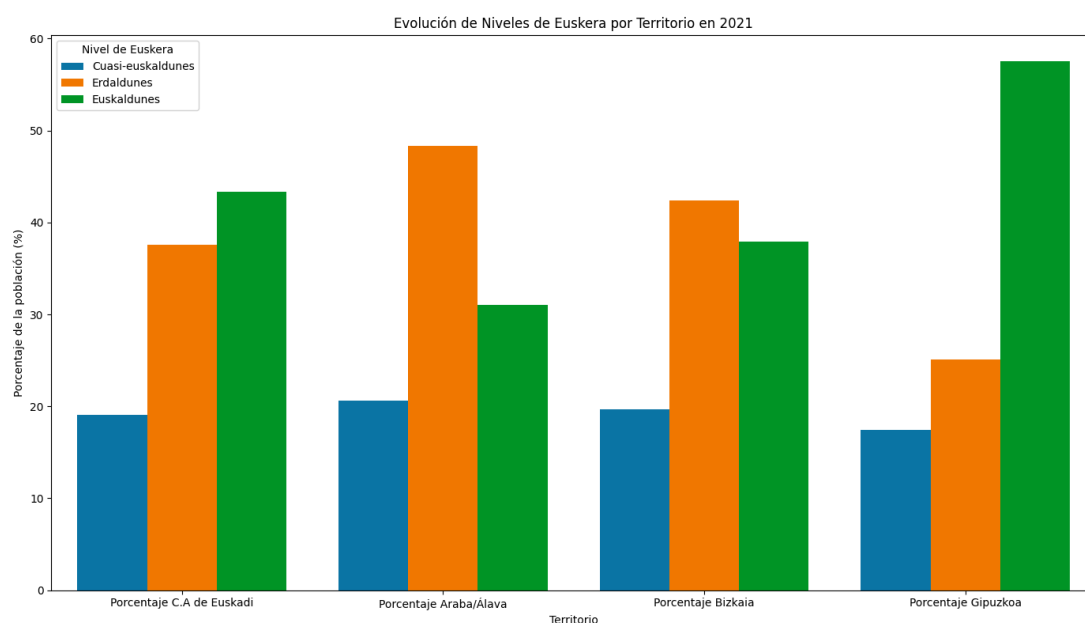
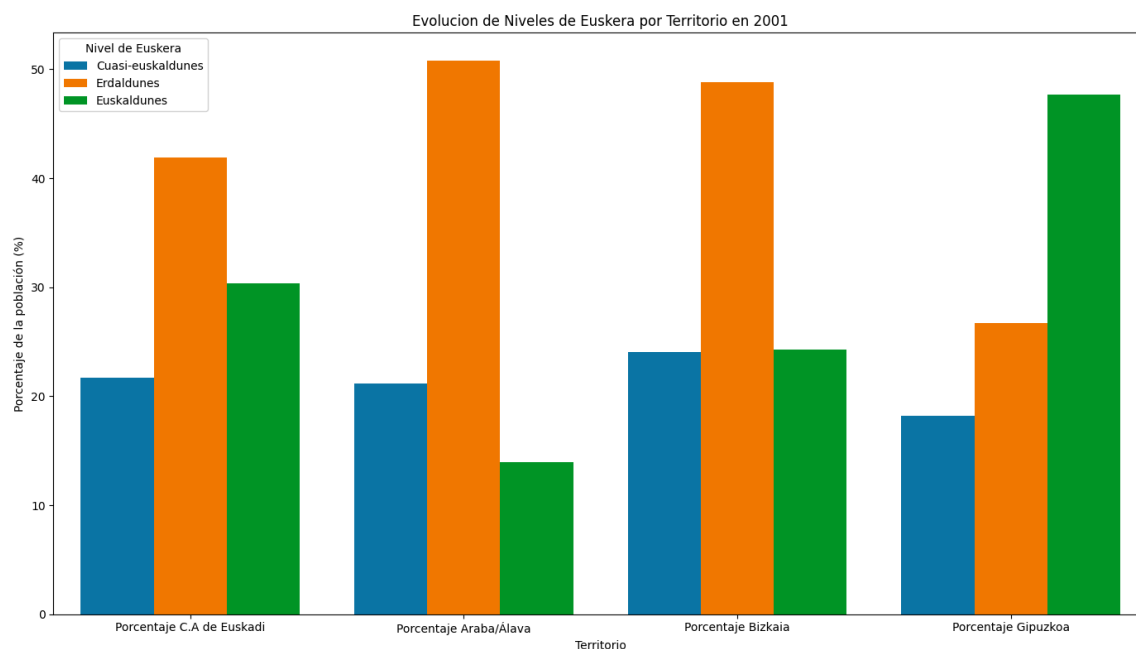
- **Euskaldunes:** personas que hablan y entienden euskera perfectamente
- **Cuasi-Euskaldunes:** personas que entienden pero tienen dificultades al hablar
- **Erdaldunes:** personas que no saben nada de euskera

5.1.1 Análisis de la evolución del conocimiento de euskera (2001-2021)



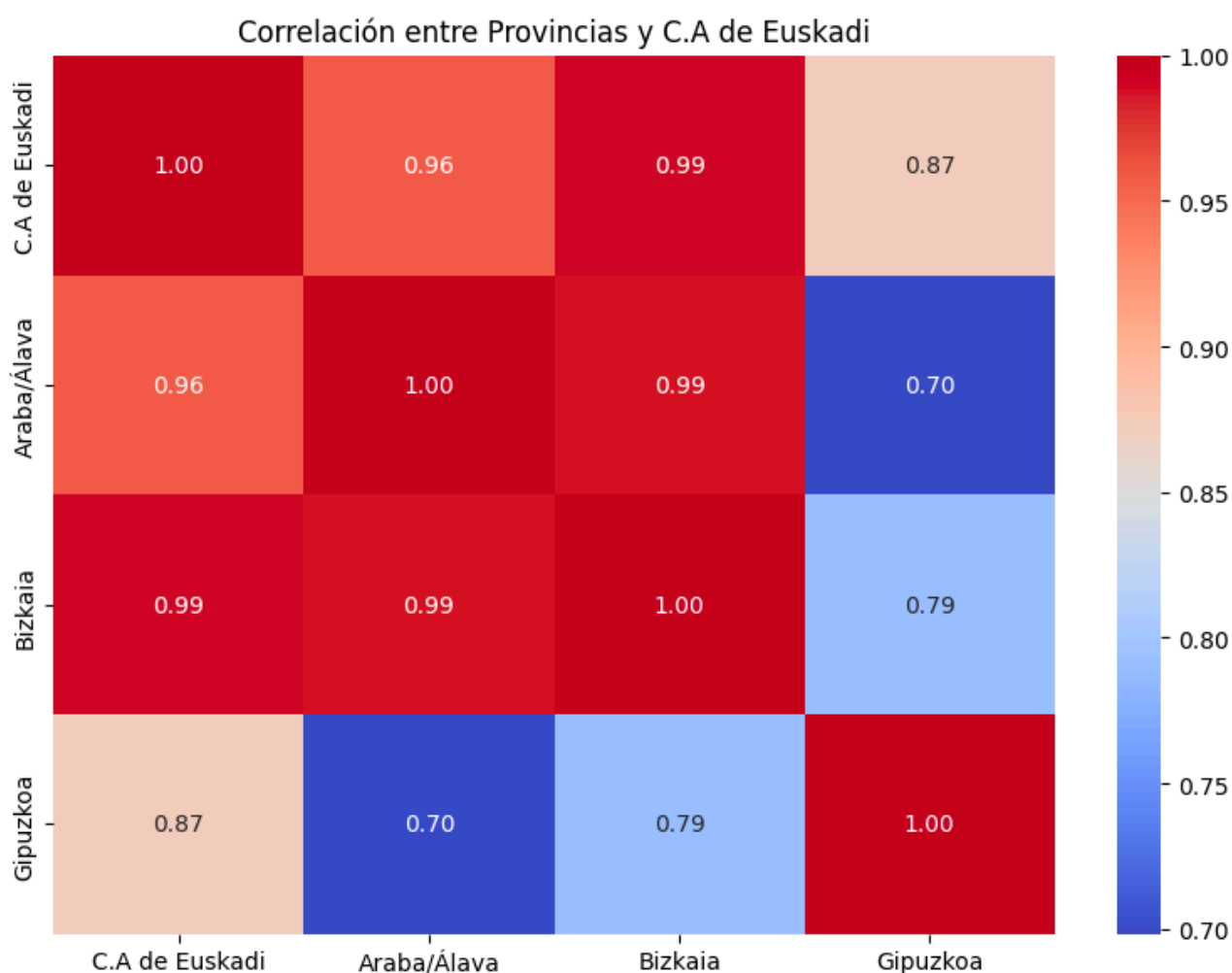
- **Aumento de la población que sabe euskera:** El número de personas que sabe euskera (Euskaldunes y Cuasi-euskaldunes) ha aumentado entre 2001 y 2021, aunque de forma modesta. El 62,4% de personas tiene algún conocimiento de euskera en 2021 respecto al 55% de personas en 2001.
- **Disminución de los Cuasi-euskaldunes:** La cantidad de Cuasi-euskaldunes ha caído, por lo que puede ser que aquellos que tenían dificultades para hablar el idioma ahora podrían ser clasificados como Euskaldunes.
- **Impacto de las políticas lingüísticas:** El incremento de hablantes de euskera puede atribuirse a las políticas lingüísticas en la educación (que desde hace pocos años la educación solo es en euskera) y la promoción del euskera.

5.1.2 Análisis de la distribución por territorio (2001-2021)



- Gipuzkoa sigue siendo la provincia con la mayor proporción de hablantes de euskera, mientras que Araba tiene una menor proporción. No sabemos a que se puede deber, pero probablemente refleja las diferencias históricas en la lengua.
- Bizkaia muestra un aumento en la población que habla euskera, pero sigue estando rezagada respecto a Gipuzkoa.
- De todos modos, Euskadi muestra que tiene una población euskadun en 2021 (62,4) respecto a 2001 que tenía una población mayoritaria erdalduna, es decir, que casi la mitad de la población (45%) no entendía euskera.

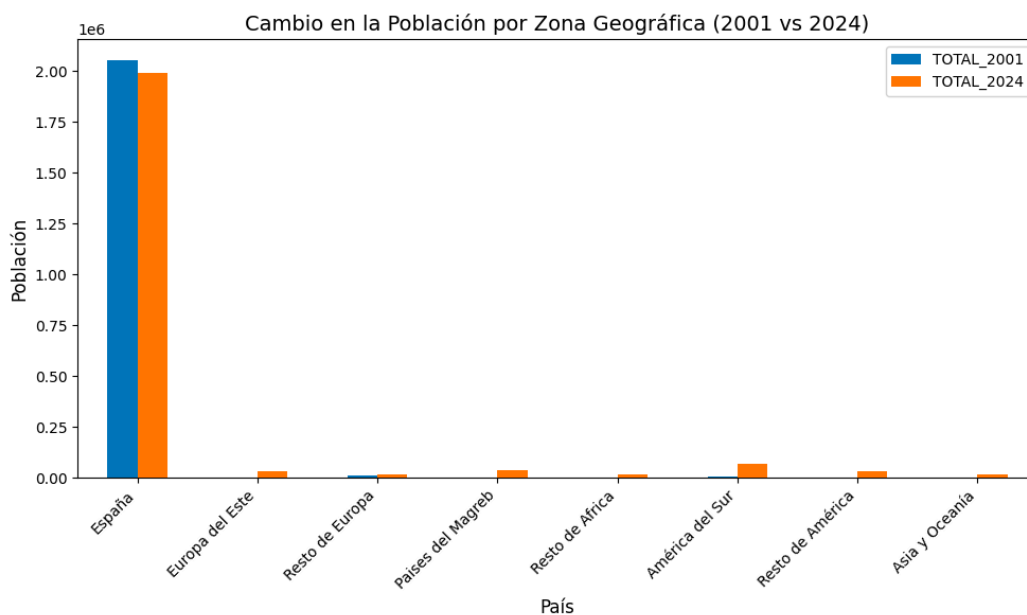
5.1.3 Análisis de la correlación entre provincias

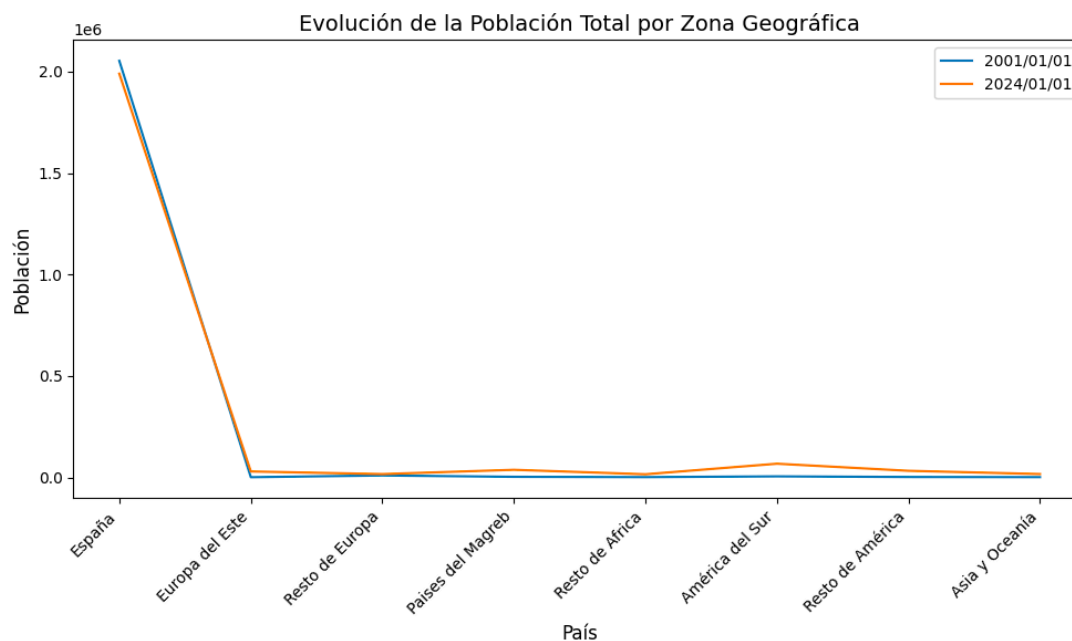
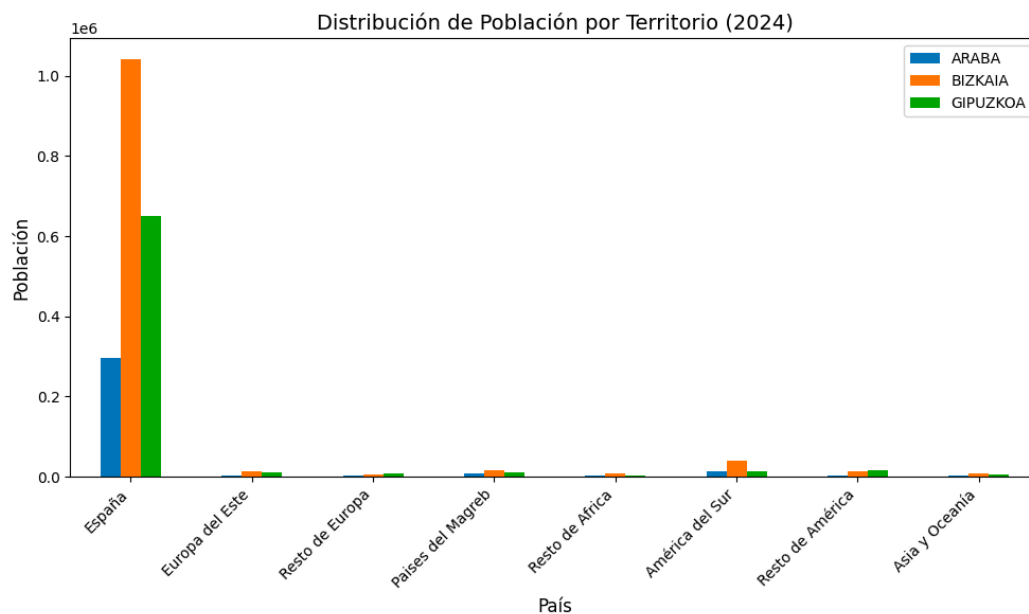


- Alta correlación entre las provincias: Tanto Álava como Bizkaia tienen una alta correlación con Gipuzkoa en cuanto al porcentaje de hablantes de euskera en 2001 y 2021. Esto podría indicar que, en general, las tendencias en el uso y conocimiento del euskera tienden a seguir patrones similares entre estas provincias. Sin embargo, Gipuzkoa se caracteriza históricamente por un uso más fuerte del euskera, por lo que las provincias como Álava y Bizkaia aún están adaptándose a este modelo.

- Desajuste en comparación con Gipuzkoa: Aunque las correlaciones son fuertes, los valores en Álava y Bizkaia todavía no alcanzan los niveles de Gipuzkoa, especialmente cuando comparamos las estadísticas entre los años 2001 y 2021. Esto puede ser un indicativo de que Gipuzkoa sigue siendo la provincia con el porcentaje más alto de hablantes de euskera, y que las otras provincias todavía tienen mucho por hacer en cuanto a la promoción y uso del idioma.
- Inmigración e influencia sociocultural: Las provincias de Álava y Bizkaia han podido experimentar un mayor flujo de inmigrantes, lo que puede haber influido en la menor adopción del euskera en comparación con Gipuzkoa. Esta diferencia en la composición demográfica podría ser una de las razones para la discrepancia en los niveles de euskaldunes, cuasi-euskaldunes y erdaldunes.
- Posibles políticas lingüísticas: Las diferencias también podrían reflejar cómo se implementan las políticas lingüísticas en las distintas provincias. Por ejemplo, Gipuzkoa ha sido históricamente más proactiva en la promoción del euskera, lo que se refleja en sus mayores tasas de hablantes del idioma.

5.2 Análisis de Población





Se exploraran las relaciones entre el conocimiento del euskera y la inmigración. Se observa que, aunque la inmigración ha aumentado en provincias como Bizkaia y Gipuzkoa en numero aunque no en proporción, la adopción del euskera sigue siendo más baja en Álava, lo que podría estar relacionado con la menor exposición al idioma en áreas con mayor diversidad lingüística.

Si los datos muestran que la población procedente de España ha disminuido en lugar de aumentado, las implicaciones son las siguientes:

Para empezar Bizkaia es la provincia mas poblada, acumulando un 52% de toda la población del Pais Vasco, seguida de Gipuzkoa con un 33% y Alava con un 15%. En 2001 la población era de 2.079.210 habitantes y en 2024 se contabilizan 2.208.007 por lo que ha la población vasca ha

aumentado un 6,19%. En 2021 la procedencia mayoritaria era de España siendo un 98,78% mientras que ahora, que sean autóctonos o de procedencia española representa el 90,15%, bajando 8 puntos que se recuperan y aumenta la población gracias a la inmigración.

Decrecimiento de la población española:

Este descenso podría reflejar emigración interna desde Euskadi hacia otras regiones de España, o un envejecimiento de la población autóctona sin un número suficiente de nacimientos para mantener la población. Podría ser también una indicación de que las nuevas generaciones no están permaneciendo en la región y se trasladan a otros lugares en busca de trabajo o mejores oportunidades.

Aun así, España sigue siendo el país dominante en términos de población en Euskadi, representando más del 90% de media la población en euskadun 2024.

Emigración interna:

En algunos casos, el decrecimiento de la población de España podría estar relacionado con el desplazamiento de jóvenes o de la población activa hacia otras regiones que ofrecen mejores oportunidades laborales, sobre todo hacia grandes centros urbanos como Madrid, Barcelona u otras zonas con mercados laborales más dinámicos.

Envejecimiento de la población:

Es probable que haya un envejecimiento de la población autóctona, donde los nacimientos no compensan la cantidad de muertes, y, como resultado, se ve una disminución de la población autóctona, mientras que los inmigrantes (por ejemplo, de Europa del Este o América del Sur) siguen siendo una parte importante del crecimiento poblacional.

El crecimiento poblacional en Euskadi probablemente esté siendo impulsado por inmigrantes internacionales, especialmente de Europa del Este, América Latina, y otras regiones, lo que implica una diversificación cultural en la región.

América del Sur tiene una representación más alta en todas las regiones. Esto indica que los flujos migratorios provenientes de países latinoamericanos han sido importantes en los últimos años, especialmente en las provincias más urbanizadas siendo Alava y Bizkaia los que mas reciban.

Europa del Este muestra un aumento moderado en su representación, especialmente en Gipuzkoa, donde la proporción alcanza el 1.6%. Esto refleja los flujos migratorios desde Europa del Este, una tendencia que ha aumentado en las últimas décadas.

Resto de África y Países del Magreb tienen una mayor representación en Álava, especialmente los países del Magreb, con un porcentaje del 2.75%. Esto refleja las dinámicas migratorias desde el norte de África hacia el País Vasco, especialmente en Álava.

Asia y Oceanía tienen una representación pequeña, pero no despreciable, con valores en torno al 0.8% en las tres provincias, lo que sugiere una migración más reciente y aún de bajo volumen desde estos territorios

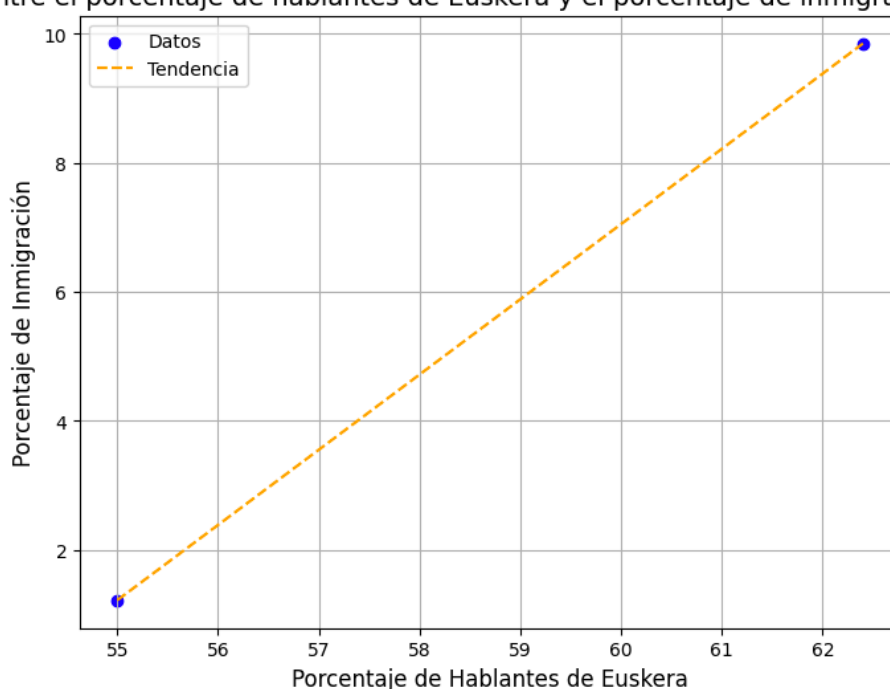
Con todo esto, hay que destacar que Álava sigue siendo la provincia con la menor población total, pero la diversidad étnica y cultural parece ser más alta en términos porcentuales. Especialmente en Álava, los grupos de origen Magrebí, Africano y Americano presentan un mayor porcentaje en comparación con las otras provincias. Esto puede ser indicativo de que Álava tiene una tasa más alta de inmigración de ciertos países o regiones.

Bizkaia y Gipuzkoa han experimentado un crecimiento en la inmigración, con un aumento en los porcentajes de países como Europa del Este (especialmente en Gipuzkoa) y América del Sur (principalmente en Álava).

Este cambio en la composición demográfica podría tener implicaciones sociales y lingüísticas, ya que la inmigración internacional podría afectar la preservación y uso del euskara.

En Álava, tal y como he indicado, la proporción de inmigrantes es más alta, lo que puede significar que este grupo demográfico no tenga tanta exposición al euskera, sobre todo si el castellano es la lengua dominante en su entorno. Esto puede influir negativamente en el aprendizaje del euskera por parte de los inmigrantes.

Relación entre el porcentaje de hablantes de Euskera y el porcentaje de inmigración (2001-2024)



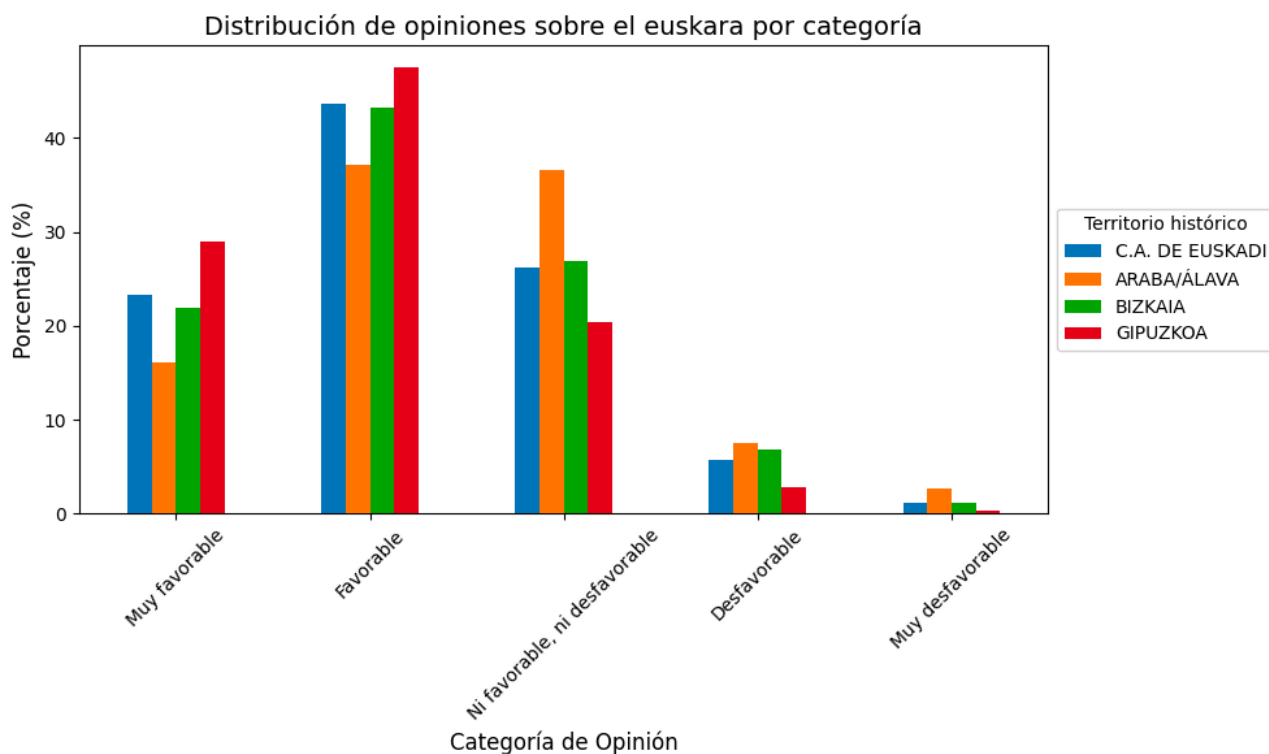
En este caso hay una correlación positiva, por lo que significa que ambas variables tienden a aumentar al mismo tiempo. Tanto el porcentaje de hablantes de euskera como el porcentaje de inmigración han aumentado entre 2001 y 2024. Esto puede ser interesante, pero no indica causalidad. Es decir, el aumento de la inmigración no "causa" directamente el aumento del conocimiento del euskera. Podrían haber otras variables influyendo en ambos factores como que el porcentaje de erdaldunes ha bajado (puede que hayan estudiado euskera y hayan pasado a los otros grupos).

No podemos fiarnos estadísticamente ya que se cuenta con muy pocos datos.

Analizamos también el porcentaje de hablantes de euskera y el porcentaje de inmigración en Álava en 2001 y 2021 y otra vez, da una relación positiva pero al haber muy pocos datos no se puede establecer una relación de confianza entre las dos variables. Hay otros factores que pueden estar influyendo en este caso. Mas bien podríamos interpretar de los datos una correlación moderada en Álava, lo que la inmigración puede estar afectando el conocimiento del euskera, especialmente entre los inmigrantes recientes.

5.3 Análisis de Promoción del Euskera

Analizamos la promoción del euskara en las diferentes regiones. Tal y como indicábamos anteriormente, Gipuzkoa es la que mas impulsa el uso del idioma respecto a las demás provincias.



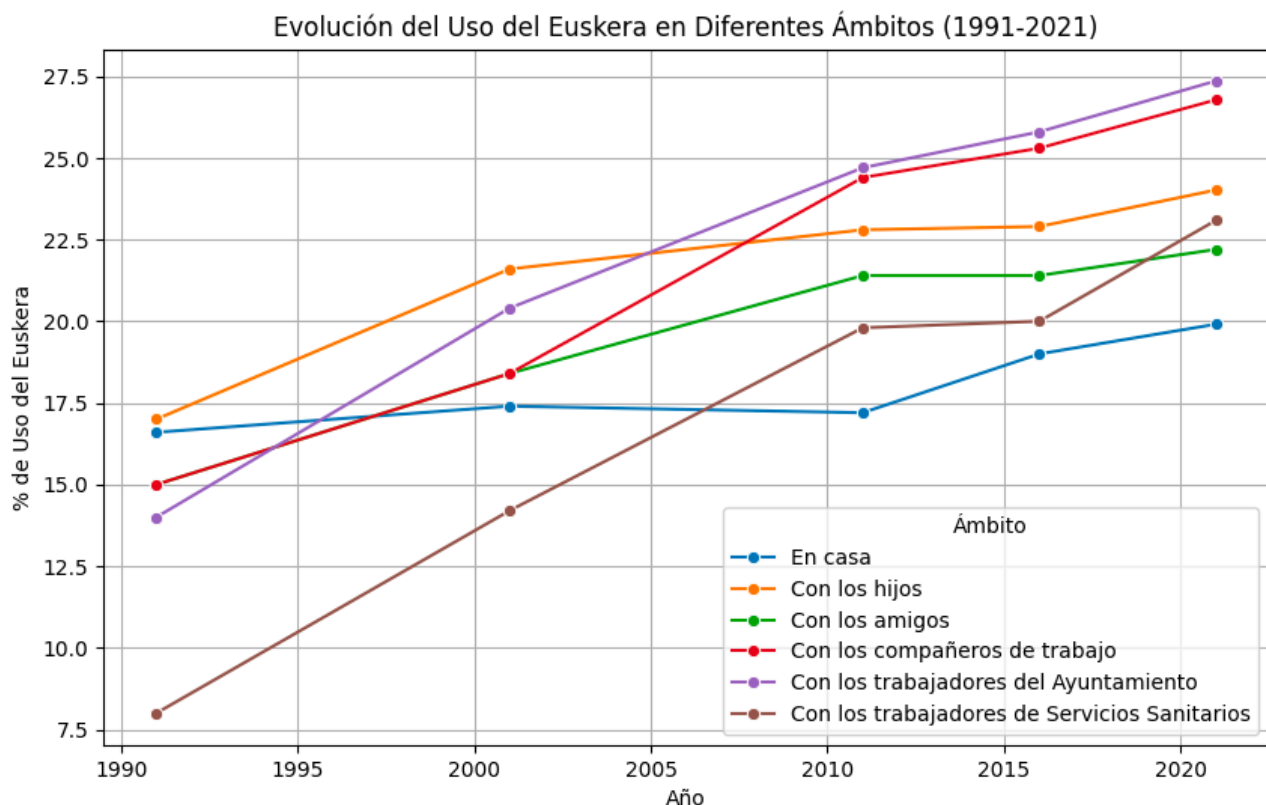
- Álava podría tener una mayor proporción de personas que se ubican en la categoría "Ni favorable, ni desfavorable" pero a su vez es el que mas destaca en posicionarse en muy desfavorable de impulsar el idioma.
- Gipuzkoa podría tener una mayor proporción de personas "Muy favorable".
- Bizkaia podría mostrar una distribución más equilibrada entre las opiniones favorables y desfavorables.

5.4 Análisis del uso del Euskera

Crecimiento general: Se observa una tendencia creciente en el uso del euskera en todos los ámbitos desde 1991 hasta 2021, lo que indica un progreso en la normalización y uso de la lengua en la vida diaria.

Ámbitos clave: Los ámbitos laborales y con los hijos muestran los mayores incrementos en el uso del euskera, lo que puede reflejar políticas públicas y educativas más eficaces.

Ámbitos con menos impacto: Aunque ha habido un crecimiento, el uso del euskera en el hogar sigue siendo moderado.



Aun así, el uso del euskera en cualquier ámbito no supera el 30%. Que mas del 50% sepa euskera en el País Vasco no significa que lo usen o lo promuevan.

6. Conclusiones

La hipótesis 1 se cumple mientras que la hipótesis 2 no, porque no se puede determinar que afecte negativamente.

Se ha observado una tendencia creciente en el conocimiento y uso del euskera en Euskadi desde 2001 hasta 2021, reflejando un avance en la normalización del idioma. Este incremento es especialmente notable en las nuevas generaciones, gracias en parte a las políticas lingüísticas implementadas por el gobierno vasco, que han promovido su enseñanza en las escuelas a través del modelo educativo D. Aunque el crecimiento es evidente, no es uniforme, ya que la adopción del euskera sigue siendo más significativa en algunas provincias (como Gipuzkoa) en comparación con otras (como Álava).

Aunque la inmigración ha aumentado en Euskadi, su impacto en el uso del euskera no parece ser decisivo. Las regiones con una mayor concentración de inmigrantes, como Álava y Bizkaia, han experimentado un crecimiento en la población, pero este fenómeno no ha conducido a un aumento significativo en el conocimiento o uso del euskera entre los inmigrantes, principalmente debido a las barreras lingüísticas y la baja exposición al idioma en su entorno social. En este sentido, la inmigración, especialmente desde países no hispanohablantes, podría estar contribuyendo a la lentitud en la adopción del euskera, aunque no de manera directa.

La correlación entre el uso del euskera y el crecimiento de la población, especialmente la inmigración, muestra una relación positiva, aunque no causal. No se pueden sacar conclusiones contundentes de que la inmigración este afectando negativamente al idioma.

En conclusión, el uso del euskera ha experimentado un crecimiento notable en los últimos 20 años, impulsado por políticas lingüísticas y educativas. Sin embargo, este crecimiento no ha sido uniforme, y el idioma sigue enfrentando desafíos significativos, particularmente en contextos familiares y en áreas con una alta diversidad lingüística. La relación entre el aumento de la población inmigrante y la adopción del euskera es compleja y requiere políticas más adaptadas a la realidad sociolingüística de Euskadi. El futuro del euskera dependerá en gran medida de la capacidad de la sociedad vasca para integrar a todos sus habitantes en un modelo de bilingüismo inclusivo y de la continuidad de esfuerzos institucionales para fomentar su uso en todos los ámbitos de la vida.

Retos y Oportunidades para el Futuro:

- **Retos:** Aunque el conocimiento de euskera ha aumentado, el uso del idioma sigue estando limitado en muchos ámbitos sociales, especialmente en el hogar y entre la población inmigrante. Las políticas públicas, aunque efectivas en algunos aspectos, deben seguir adaptándose para fomentar la adopción del euskera entre las nuevas generaciones y especialmente entre los inmigrantes.
- **Oportunidades:** Existe una gran oportunidad para fortalecer la enseñanza del euskera en la infancia y entre los adultos, particularmente a través de programas de inmersión que incluyan a la población inmigrante. Además, las políticas de normalización lingüística deben ser más inclusivas, asegurando que el euskera no sea solo un idioma de enseñanza, sino también un idioma presente en todos los aspectos de la vida social y laboral.

7. Bibliografía

Eustat: <https://www.eustat.eus/indice.html>