

Introducción

El desarrollo socioeconómico de los países está determinado por una interacción compleja de factores. Este estudio se enfoca en identificar y comparar los factores más influyentes sobre las tendencias de crecimiento de Argentina, utilizando como referencia países con desarrollos similares tanto en Latinoamérica como en otras regiones seleccionadas.

Con una perspectiva práctica, buscamos proporcionar datos relevantes que puedan servir como referencia en la elaboración de estudios futuros. Nos centraremos en comprender: ¿Qué indicadores que dan cuenta de factores socioeconómicos han influido significativamente en el desarrollo de Argentina durante las últimas seis décadas y cuál es su correlato en países con desarrollos comparables?

Metodos y Materiales

Datos

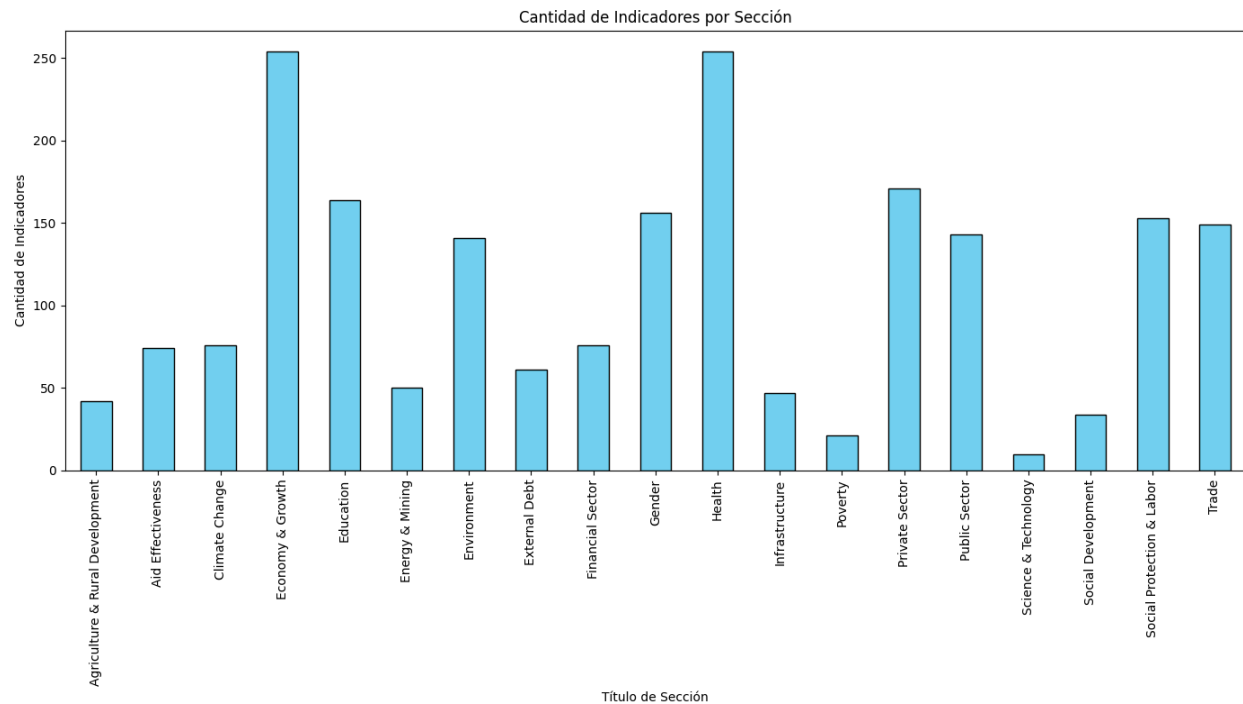
Trabajaremos con un dataset del Banco Mundial que incluye alrededor de 1462 indicadores de países en áreas clave como Educación, Ciencia y Tecnología, y Crecimiento económico, abarcando el período de 1960 a 2021. Los datasets se encuentran disponibles en este [link](#) de descarga.

Preprocesamiento

Los datos se encuentran agrupados en distintas planillas por cada indicador. El trabajo de preprocesamiento consiste en los siguientes pasos:

1. Descarga automatica: Cada indicador se encuentra en un archivo excel separado del resto.
2. Generacion de metadata: Ademas de bajar la informacion, es importante generar un archivo maestro con la metadata necesaria para el procesamiento y entendimiento de la informacion (grupos de indicadores, significado de cada indicador, etc.)
3. Agrupamiento: Luego de descargar los archivos, y para facilitar su uso se genero un proceso que agrupa los datos en las categorias anteriormente descriptas. Esto permite, en lugar de tener 1462 archivos de datos separados, tener 19 archivos.

En la siguiente figura pueden verse las cantidades de indicadores por cada grupo.



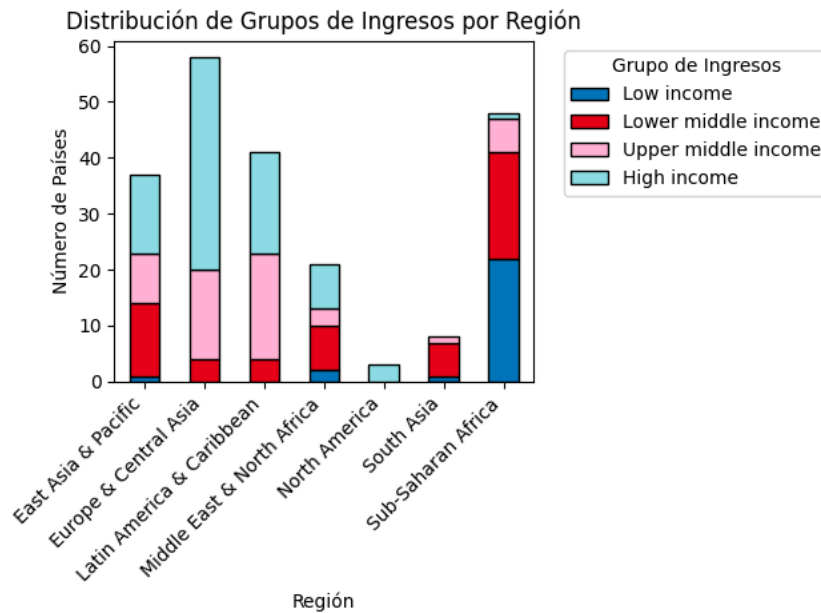
Es importante mencionar que los distintos países ya se encuentran agrupados por región y categoría de income.

Regiones:

- 'Latin America & Caribbean'
- 'South Asia'
- 'Sub-Saharan Africa',
- 'Europe & Central Asia',
- 'Middle East & North Africa',
- 'East Asia & Pacific',
- 'North America'

Categorías de income:

- High income group aggregate. High-income economies are those in which 2022 GNI per capita was more than \$13,845.
- Low income group aggregate. Low-income economies are those in which 2022 GNI per capita was \$1,135 or less.
- Lower middle income group aggregate. Lower-middle-income economies are those in which 2022 GNI per capita was between \$1,136 and \$4,465.
- Low and middle-income group aggregate. Low and middle-income economies are those in which 2022 GNI per capita was less than \$13,845.



Analisis y Modelado

Nuestro análisis constará de: (1) Métodos descriptivos que incluyen técnicas estadísticas y visualización de datos para caracterizar y comparar países; (2) Técnicas de clustering como K-means para identificar agrupaciones de países con patrones de desarrollo similares; y (3) Análisis de series temporales utilizando modelos como ARIMA para evaluar la evolución de los indicadores a lo largo del tiempo.

Métodos descriptivos

El foco estará dado en entender si los datos existentes permiten la utilización de las técnicas de clustering y de análisis de series temporales que pensamos utilizar. En particular, existen muchos datos faltantes, que pueden dificultar la tarea. En estos casos intentaremos no imputar valores sino utilizar los indicadores que se encuentren más completos.

Los datos son confiables pero hay muchos faltantes. Voy a tener que acotar los indicadores a aquellos que estén más completos y tal vez a periodos de tiempo más acotados y a regiones o categorías de países que se encuentren más completos para poder hacer clustering.

Técnicas de clustering

Análisis de series temporales

Herramientas

Resultado esperado

Esperamos que este enfoque analítico revele patrones subyacentes en la trayectoria de desarrollo de los países estudiados, ofreciendo así una base de conocimiento que pueda servir para futuras investigaciones y para la formulación de políticas más efectivas hacia el desarrollo sostenible.