

Análisis exploratorio de datos (EDA)*

JOSÉ ÁNGEL TORRENS HERNÁNDEZ, 175021 | Ciencia Política

Introducción

Se presenta un análisis exploratorio de datos (EDA) no exhaustivo para algunas variables de las bases de datos *Human Development Index (hdi . csv)* del Programa de las Naciones Unidas para el Desarrollo UNDP [2024], y el conjunto de datos *World Income Inequality Database (wiidglobal . csv)* del Instituto Mundial de Investigación sobre Economía del Desarrollo de la Universidad de las Naciones Unidas UNU-WIDER [2023].

Estadísticas descriptivas

Los dataset utilizados contienen información sobre desarrollo humano, pobreza y desigualdad para más de 190 países y territorios. A continuación se presentan algunas estadísticas descriptivas de las variables de interés.¹

Table 1: Estadísticas Descriptivas (Serie 1950 - 2022)

Statistic	N	Mean	St. Dev.	Min	Max
population	15,257	24,140,983.0	100,784,269.0	1,458	1,425,893,504
gdp	15,257	14,200.6	19,829.4	292.9	215,779.5
gini	15,257	43.8	11.6	0.0	74.2
bottom5	15,257	0.8	0.5	0.05	3.1
bottom20	15,257	5.3	2.4	0.2	14.0
bottom40	15,257	14.9	5.2	0.4	30.6
top5	15,257	23.3	8.1	0.1	51.8
top10	15,257	34.5	9.6	0.1	65.4
top20	15,257	39.4	5.1	0.2	52.3
middle50	15,257	50.6	5.9	29.2	99.5
abr	5,568	59.1	47.3	1.5	214.5
gdi	4,360	0.9	0.1	0.5	1.1
gii	4,389	0.4	0.2	0.01	0.8
hdi	4,970	0.7	0.2	0.2	1.0
le	5,568	68.5	9.9	14.1	86.5
mmr	5,216	239.0	357.7	2.0	2,990.0
pr_f	5,160	16.4	11.1	0.0	57.5
pr_m	5,160	83.6	11.1	42.5	100.0

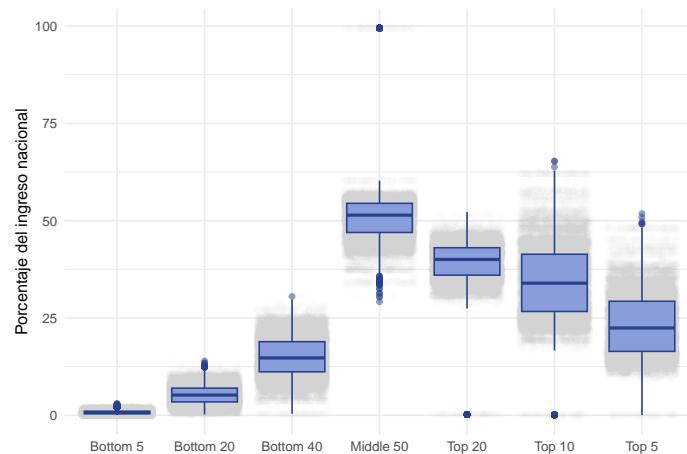
* Los archivos del dataset se encuentran disponibles en (<https://github.com/pptrrns/infoViz/tree/main/dataset>).

¹ Población (population); ingreso per cápita (gdp); Coeficiente Gini (gini); % del ingreso nacional del 5% peor remunerado (bottom5), del 20% peor remunerado (bottom20), del 40% peor remunerado (bottom40), del 5% mejor remunerado (top5), del 10% mejor remunerado (top10), del 20% mejor remunerado (top20) y del 50% de la población (middle50); tasa de nacimientos adolescentes (abr); Índice de Desarrollo de Género (gdi); Índice de Desigualdad de Género (gii); Índice de Desarrollo Humano (hdi); esperanza de vida (le); tasa de mortalidad materna (mmr); años promedio de escolarización (mys); % de asientos en el Parlamento ocupados por mujeres (pr_f) y % ocupados por hombres (pr_m).

Visualizaciones

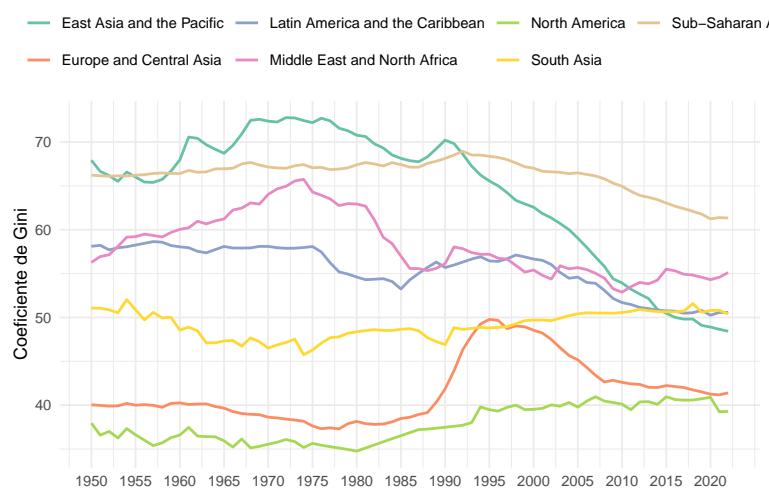
Desigualdad por ingreso

Se observa la distribución del porcentaje de participación del ingreso nacional entre diferentes grupos de la población. o anterior permite analizar la desigualdad en la distribución del ingreso entre diferentes percentiles. Del gráfico podemos observar que, a nivel global, el *Top 5* de las personas mejor remuneradas concentra el 23.3% del ingreso (promedio entre países).



Coeficiente de Gini a nivel global, regional y nacional (Serie 1950 - 2020)

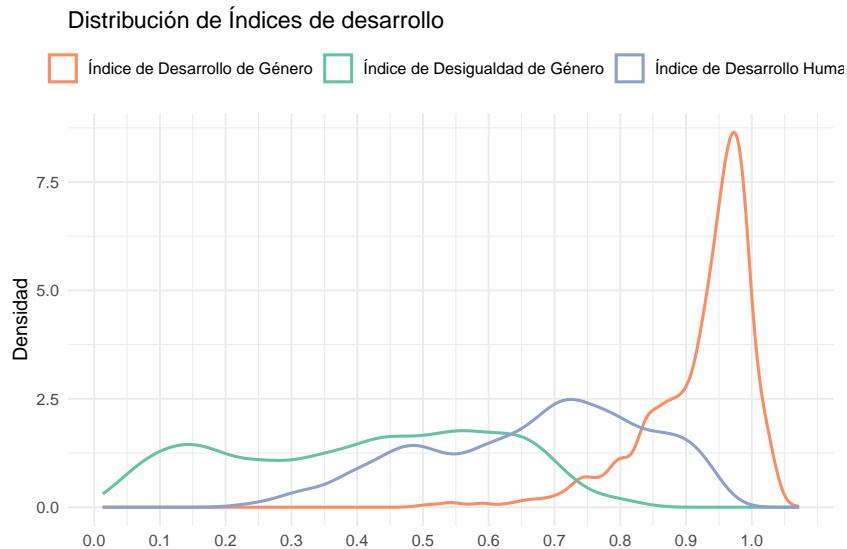
A continuación se presentan la evolución del Coeficiente de Gini a nivel regional. El Coeficiente es una medida para estimar la desigualdad en los ingresos dentro de un país y se computa como una variable entre 0 y 1, donde 0 se corresponde con la perfecta igualdad y 1 se corresponde con la perfecta desigualdad.²



²Para el entregable final el usuario podrá seleccionar el país, hacer comparaciones entre países, regiones y los datos a nivel global. Además podrá desplegar la visualización para algunos grupos de países, como el G7, G20, la Organización para la Cooperación y el Desarrollo Económicos y los BRICS. Además, el entregable final permitirá realizar las mismas visualizaciones para indicadores como el Índice de Desarrollo Humano, la esperanza de vida al nacer, la tasa de mortalidad materna, y los años promedio de escolarización.

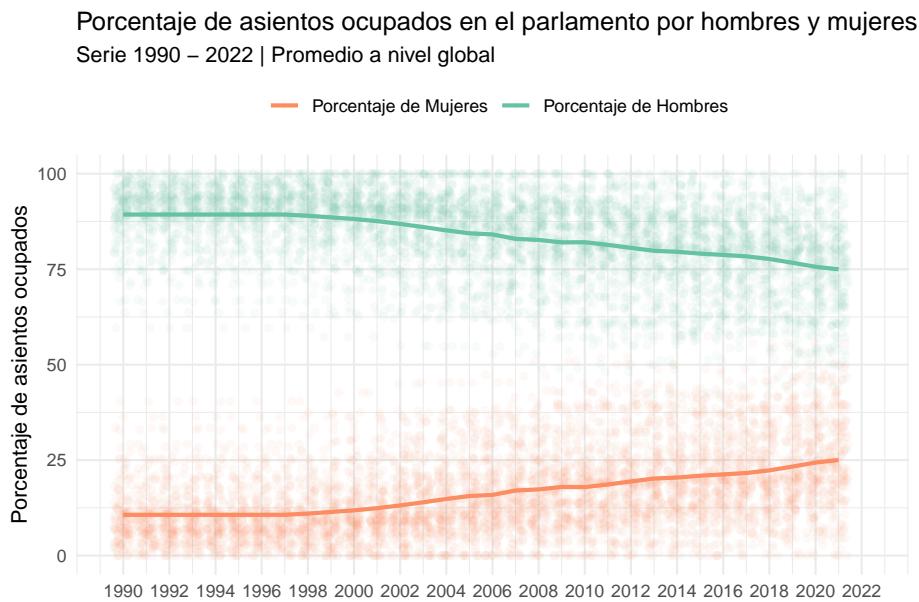
Desigualdad por género

Se presenta la distribución de algunos indicadores de interés, como el Índice de Desarrollo de Género (gdi) y el Índice de Desigualdad de Género (gi). Los datos nos muestran cómo se distribuye la evaluación de más de 190 países en estos indicadores.



Desigualdad: participación política

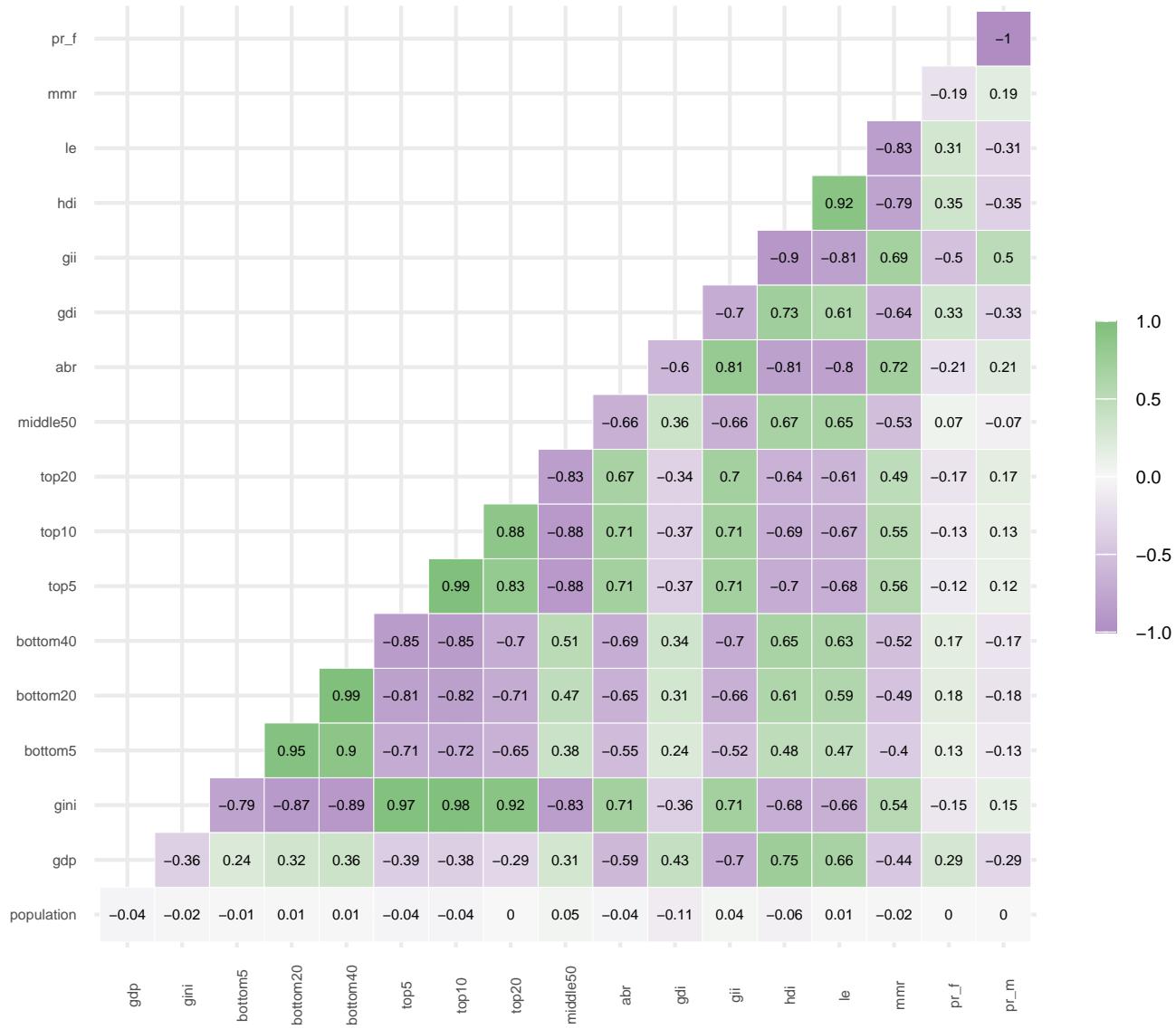
Se presenta la serie histórica sobre el porcentaje de asientos ocupados en el Parlamento por hombres y mujeres. La línea muestra la tendencia a nivel global (promedio de países), mientras que el marcador de los puntos las observaciones a nivel país. Con los datos se puede apreciar una ligera disminución del porcentaje de asientos ocupados por hombres, así como un aumento en los ocupados por mujeres desde 1990 hasta 2020.³



³Para el entregable final, el usuario podrá interactuar con las obs. a nivel país y hacer comparaciones entre países y regiones.

Matriz de Correlación

Se presenta una matriz de correlación entre las variables de interés para el proyecto final.



References

UNDP. Human development index (hdi). Technical report, United Nations Development Programme, 2024.

UNU-WIDER. World income inequality database (wiid). Technical report, United Nations University World Institute for Development Economics Research (UNU-WIDER), 2023.