

2. Titulo

Dataset: Ranking Mundial de la Felicidad 2021

1. Contexto

El "Ranking Mundial de la Felicidad" consiste en obtener una calificación(índice) y ranking por país por medio de una encuesta que se aplica a la población de los países participantes. De acuerdo con la calificación obtenida por cada país, se obtiene un ranking a nivel global.

Esta iniciativa tiene por objetivo medir diferentes factores sociales, salud, económicas y seguridad en la población para usar estos datos en la generación de políticas públicas que beneficien a la sociedad.

Link del sito utilizada para extraer datos:

<https://es.theglobaleconomy.com/rankings/happiness/>

3. Descripción del dataset

Conjunto de datos que muestra la calificación(índice) y ranking que ocupa cada país respecto al índice de felicidad.

4. Representación gráfica

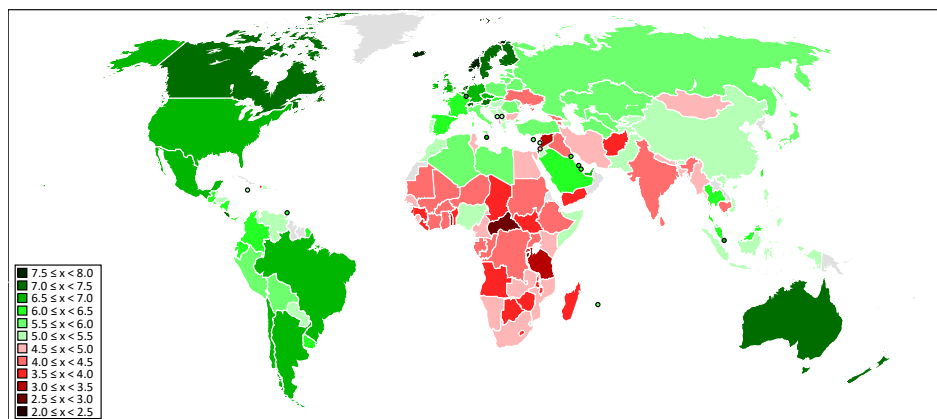


Figura 1. Calificación(Índice) de los países participantes en el ranking mundial de la felicidad 2018. En la escala de color, los países con una tonalidad verde-oscuro son los más felices mientras que la tonalidad rojo-oscuro son los que tienen más oportunidad de desarrollar políticas sociales para mejorar su calificación.

5. Contenido

Cada país participante en el ranking mundial de la felicidad, se obtienen las siguientes características:

- Country_name: Nombre del país.
- Index_of_happiness: Calificación obtenida.
- Global_rank: Posicionamiento del país a nivel mundial.
- Available_data_date: Periodos disponibles.
- Region: Región al que pertenecen los países.

6. Propietario

TheGlobalEconomy.com es la fuente de información, este sitio tiene como propósito servir a investigadores, empresarios, estudiantes y académicos quienes necesiten datos socio-económicos de los diferentes países del mundo. Por citar algunas fuentes que este sitio utiliza son:

- [The World Bank](http://TheWorldBank.org)
- [The United Nations](http://TheUnitedNations.org)
- [Transparency International](http://TransparencyInternational.org).

Este estudio se ha estado realizando desde el año 2013.

El presente proyecto tiene **finances netamente académicos, sin fines comerciales**, por lo cual no entra en conflicto con las políticas que maneja el sitio TheGlobalEconomy.com ([Privacy policy and Economic data feed via XML API](#))

7. Inspiración

Como se mencionó anteriormente, el propósito del ranking mundial de la felicidad es medir diferentes factores sociales, salud, seguridad y económicos de la población de los países participantes y así tener una calificación(índice) y ranking. Tenemos 2 enfoques para este estudio:

- Países dentro de los primeros lugares en el ranking: Se recomienda al menos continuar y en la medida de lo posible mejorar con las políticas sociales, salud, económicas y de seguridad para que la población siga estando feliz.
- Países que no están dentro de los primeros lugares en el ranking: Se recomienda mejorar las políticas sociales, salud, económicas y de seguridad para que la población se sienta más feliz.

Medir el factor de la felicidad se está convirtiendo en un eje muy importante en el mundo para el desarrollo de los países. Estudiar el hecho de que la gente no este feliz puede llevar al gobierno en turno, a la aplicación de diferentes políticas que puedan apoyar a la sociedad a sentirse mejor y por consecuencia, a mejorar el ámbito laboral y personal.

8. Licencia

El dataset generado tendrá la licencia, Attribution-NonCommercial- ShareAlike 4.0 International (CC BY-NC-SA 4.0), los motivos son los siguientes:

- Attribution: Se debe de dar crédito a la persona que creo el conjunto de datos e indicar el nombre de la licencia.
- NonCommercial: Sin fines comerciales.
- ShareAlike: Cualquier modificación que se haga posterior a la versión original, tendrá que ser distribuido con el mismo licenciamiento al original.

9. Código

Aspectos relevantes del código:

- Se utiliza la librería **selenium** para las tareas de scraping.
- Se obtiene el código fuente de la página web para ser renderizado por los métodos que ofrece selenium .
- Se buscan los controles html que tienen el grid con los datos y también el combobox para poder cambiar los valores y obtener toda la historia. **Se destaca que todo esto se realiza de forma automática con las herramientas que ofrece selenium.**
- Se utilizan los métodos "find_element" con la propiedad "value" y "By.XPATH" para poder encontrar los controles html.
- Se utiliza el método "click()" que ofrece el método "find_element" para poder cambiar los datos del sitio de internet.
- Se utiliza la librería "csv " para poder exportar los datos a un archivo "csv"

Las principales dificultades para obtener los datos fueron:

- El sitio, después de hacer varias peticiones a causa de las pruebas que estaba haciendo con el código, comenzó a enviar mensajes de sitio no accesible. Se comenzó a utilizar la librería "time" y la funcionalidad "sleep" para poder hacer pausas al momento de extraer los datos históricos.
- Para mostrar los resultados de los rankings y calificaciones(índice) de un año, se utiliza un combo, el cual tiene el usuario que hacer click y poder así obtener datos de otros años, para solventar esto se utilizó los métodos mencionados anteriormente find_element() con la propiedad "value" y "By.XPATH".

10. Dataset

Target URL:	https://doi.org/10.5281/zenodo.7344818
DOI	10.5281/zenodo.7344818
Image URL	https://zenodo.org/badge/DOI/10.5281/zenodo.7344818.svg
Markdown	[![DOI](https://zenodo.org/badge/DOI/10.5281/zenodo.7344818.svg)](https://doi.org/10.5281/zenodo.7344818)
HTML	

10. Video

<https://drive.google.com/drive/folders/1JfebgnKTGclEpOZ4aioj7LF88kP8dVpC?usp=sharing>

GIT link:

https://github.com/LuisOlascoagaMX/web_scraping_ranking_happiness.git

Contribuciones	Firma
Investigación previa	LEOD
Redacción de las respuestas	LEOD
Desarrollo de código	LEOD
Participación en el video	LEOD