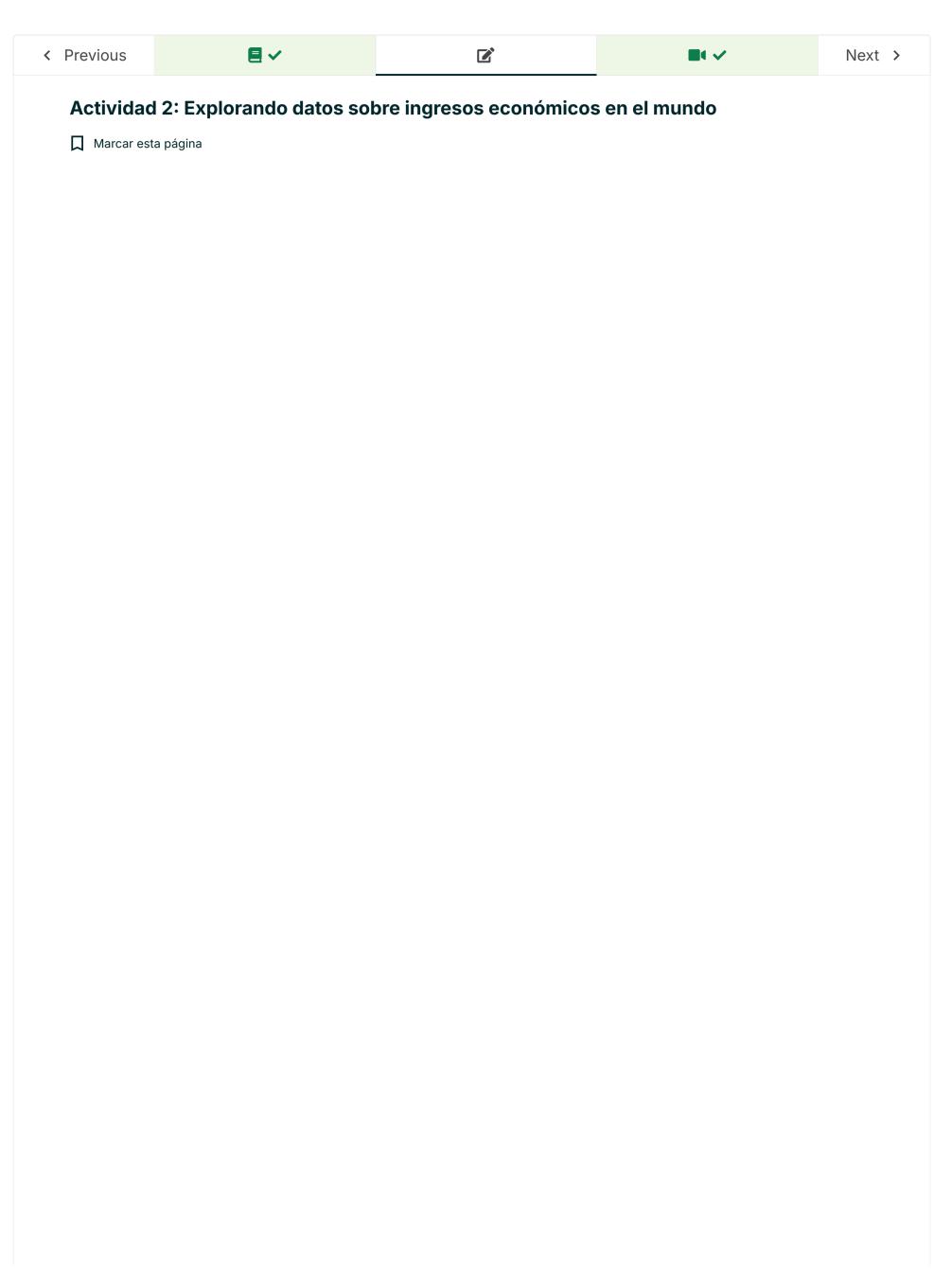


Ayuda 🗘 😢

<u>Course</u> <u>Progress</u> <u>Fechas</u> <u>Discussion</u>

☆ Curso / 2. Recolección y preparación de datos / Checkpoint #2



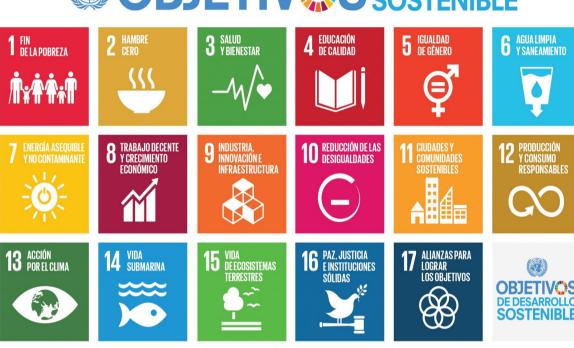
Checkpoint #2 due May 8, 2025 09:33 CST

#### Actividad 2: Explorando datos sobre ingresos económicos en el mundo



La pobreza es un tema de suma relevancia para la agenda de todos los gobiernos alrededor del mundo. Los esfuerzos de múltiples naciones se han consolidado en los Objetivos de Desarrollo Sostenible propuestos por la ONU, organización de la que México y muchos países de latinoamérica forman parte. Conocer la magnitud del problema es prioridad en orden para proponer intervenciones eficaces que abonen a disminuir la brecha de desigualdad.





Producido en colaboración con **TROLLBÄCK + COMPANY** | The Global Goals @trollback.com | +1.212.529.101

En la siguiente actividad limpiarás y explorarás un dataframe en el que encontrarás datos relevantes al análisis de la pobreza en los diferentes países que integran a la Organización de las Naciones Unidas. Al finalizar deberás subir tu cuaderno en un archivo .zip.

Te recomendamos descargar la siguiente libreta de Jupyter que preparamos para ti. Te ayudará mucho como guía en tu actividad.

#### **DESCARGA AQUI**

Es muy importante que subas tus archivos de la manera en cómo te indicamos para que tus compañeros puedan revisar tu código.

- 1. Importa las librerías de Pandas, Numpy, Requests, Webbrowser y Beautiful soup a tu libreta. Las estaremos utilizando a lo largo del ejercicio.
- 2. Accede a la siguiente página web desde tu libreta utilizando webbrowser. https://en.wikipedia.org/wiki/List\_of\_countries\_by\_percentage\_of\_population\_living\_in\_poverty
- 3. Descarga en tu libreta la página proporcionada y obtén todas las tablas utilizando la librería requests y beautiful soup. Recuerda escribir el número de tabla correcto. Te recomendamos explorar la página.
- 4. Utilizando la función "read\_html" de la librería "Pandas" guarda la tabla "Percent of population (including noncitizens) living under \$1,90, \$3,20 and \$5,50 a day" en una variable llamada "df".

  Sugerencia: Ocupa la función "head" de la librería "Pandas" para mostrar los primeros cinco elementos y verificar que haya funcionado lo anterior.
- 5. ¿Cuántas filas tiene el dataframe?

Recuerda utilizar tu libreta de Python para encontrar la respuesta

_	
30 . ¿Cuántas Recuerda	oo ns columnas tiene el dataframe?
. ¿Cuánta: Recuerda	s columnas tiene el dataframe?
Recuerda	
Recuerda	
Recuerda	
<u> </u>	
	96
7	
<b>30</b>	00
	a que los datos de la tabla contienen registros vacíos y el símbolo de porcentaje, es necesario que realices ieza de estos datos; utiliza las funciones "replace" y "dropna" de "Pandas" para lograrlo.
	de sintaxis de la función "replace": df = df.replace({'%':""}, regex = True). (Se sugiere utilizar la función con la tabla modificada para visualizar que todo se encuentre en orden).
¿Cuánta	ns filas quedaron después de hacer esto?
Recuerda	utilizar tu libreta de Python para encontrar la respuesta
6	
16	66
<b>13</b>	30
más info	el tipo de datos que poseen las columnas de tu dataframe utilizando la función "info". (Puedes encontrar ormación acerca de cómo utilizar esta función a través del siguiente link: <a href="https://www.geeksforgeeks.org/pandas-dataframe-info/">https://www.geeksforgeeks.org/pandas-dataframe-info/</a>
¿Cuál es	s el tipo de datos que aparece en la mayoría de las columnas?
Recuerda	utilizar tu libreta de Python para encontrar la respuesta
O flo	pat64
ok	oject
) int	t

Transforma las columnas "< \$1.90[1]", "< \$3.20[5]", "< \$5.50[6]" a **tipo flotante** para que puedas hacer análisis cor ellas. Para esto te recomendamos utilizar la función **"astype"**. Para más información te recomendamos entrar a el siguiente vínculo:

<u>https://www.geeksforgeeks.org/python-pandas-dataframe-astype/</u> #:~:text=astype()%20method%20is%20used,type%20to%20another%20data%20type.

10.

Utiliza la función describe en el dataframe y responde:

	rda utilizar tu libreta de Python para encontrar la respuesta
	13.98%
	38.73%
	26.21%
	2013.8%
_	l es el porcentaje máximo de personas que viven con menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta
	80.71%
	91%
	97.7%
	19.52%
	Montenegro
	Montenegro
	Congo
	Congo México
	Congo
¿Cuál Recuer	Congo  México  South Sudan  Les el porcentaje de la población en México que gana menos de \$1,90 dólares al día?  rda utilizar tu libreta de Python para encontrar la respuesta
¿Cuál Recuer	Congo  México  South Sudan  l es el porcentaje de la población en México que gana menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta  97.3%
¿Cuál Recuer	Congo  México  South Sudan  I es el porcentaje de la población en México que gana menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta  97.3%
¿Cuál Recuer	Congo  México  South Sudan  Les el porcentaje de la población en México que gana menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta  97.3%  23%  3.1%
¿Cuál	Congo  México  South Sudan  I es el porcentaje de la población en México que gana menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta  97.3%
¿Cuál	Congo  México  South Sudan  les el porcentaje de la población en México que gana menos de \$1,90 dólares al día? rda utilizar tu libreta de Python para encontrar la respuesta  97.3%  23%  3.1%  6.6%  les son los 5 países que tienen el mayor porcentaje de población que gana menos de \$1,90 dólares al día

		Montenegro, Madagascar, Cyprus, Congo, Czech Republic
15. l	Utiliza	la función "drop" de "Pandas" para eliminar la columna "year" y así simplificar más tu conjunto de datos.
16.		
		la función "groupby" y contesta la siguiente pregunta:
		es el porcentaje promedio de la población en Sudamérica que vive con menos de \$3,20 dólares al día?
F	Recuer	da utilizar tu libreta de Python para encontrar la respuesta
		9.68%
		5.6%
		13.95%
		25.45%
		es el país que está en el lugar 140 de la lista? da utilizar tu libreta de Python para encontrar la respuesta
		Switzerland
		Philippines
		Syria
		Surinam
l		
		es el país que posee el 140 como número de índice? da utilizar tu libreta de Python para encontrar la respuesta
		Switzerland
		Philippines
		Peru
		Sweden
		es el país de América del sur que tiene menos porcentaje de población ganando \$1,90 dólares al día? da utilizar tu libreta de Python para encontrar la respuesta
		Chile
		Argentina
		Colombia
		Uruguay
l		

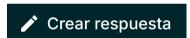
	Bolivia	
	Venezuela	
	Suriname	
	Guyana	
	l es el país de América del norte que tiene más porcentaje de población ganando \$5,50 rda utilizar tu libreta de Python para encontrar la respuesta	dólares al día?
	Haiti	
	Belize	
	Guatemala	
	Honduras	
Ejem 24. Guar	Costa Rica  Costa Rica  a la función "to_csv" de "Pandas" para generar un archivo .csv con la información obtenico de la sintaxis de la función anterior: df.to_csv("NOMBRE.csv")  da tu libreta, tu archivo csv recién creado y colócalos en una carpeta comprimida con exteno .zip a la plataforma de EdX y califica los ejercicios de tus demás compañeros.  You have used 0 of 2 attempts	
	a semana 2 es de la semana / Actividad 2: Explorando datos sobre ingresos económicos	Ocultar Discusión
Show all pos		Añade una publicación
? activida	d 2 - información wikipedia actualizada no corresponde vidad dos nos dirigen a un enlace de wikipedia el cual en las tablas ya no contiene la misma información que en el 2019	by recent activity V
<b>●</b> Evalua	ión Actividad 2  que pueda evaluar la actividad 2 por favor	1
● Hola!	evaluado mi ACTIVIDAD#1 ¿es posible que sea avaluada or el equipo del curso?	1
No se p	uede responder el ejercicio compartir la información correcta para el ejercicio?	1
Pueden		

-	Hola, La liga compartida no coincide con lo que solicitan, ¿podrían compartir la liga que hace referencia al ejercicio?	2
2	ejercicio de semana 2 Las tablas en wikipedia están actualizadas y no coinciden, eso complicó un poco.	2
<b>Q</b>	Solicitud: Repetir la Actividad_2 por Falta de Actualización  Debido a que la base de la información de la actividad_2 de la página Wikipedia está desactualizada, no es la misma a la plantaeada cuando se r	1
Q	Subir Archivo Actividad_2  Por error me falto subir el archivo de la carpeta *ipynb podrían abrirme la oportunidad para subirlo. PepeChuy2020	1
<b>Q</b>	Prueba & Base de Datos Desactualizada  La información de Wikipedia está desactualizada. Ejecuté el ejercicio con el link de la Actividad_2 y la mayoría de las respuestas no coinciden.	1
2	Sobre el enlace de wikipedia del ejercicio  Hola, después de indagar por el enlace de wikipedia que viene en el enunciado, he visto que la tabla a la que se refiere el ejercicio está archivad	5
?	Información desactualizada  Los datos han cambiado desde que se construyó el ejercicio, cómo se puede obtener la información original del ejercicio?	1
?	Desactualización información consulta  Buenas tardes, Conteste el cuestionario de forma aleatorea, debido que la información que se consulta es muy diferente a la referencxiada, por	1
Q	Interesante introducción  Me ha parecido interesante, una introducción de los conceptos estadísticos con base a python	2

### Peer to peer: Actividad 2



**Instrucciones:** Ingrese su respuesta al mensaje. Tu trabajo se guardará automáticamente y podrás regresar para completar tu respuesta en cualquier momento antes de la fecha límite. Después de enviar e respuesta, no podrá editar.



### Crear una respuesta al siguiente mensaje

Guarda tu libreta(.iypnb), tu archivo .csv recién creado y colócalos en una carpeta comprimida con extensión .zip. Sube el archivo .zip a la plataforma de EdX y califica los ejercicios de tus demás compañeros.

< Previous

Next >

© All Rights Reserved



## edX

Acerca de

<u>Afiliados</u>

edX para negocios

Open edX

Carreras

**Noticias** 

# Legal

Condiciones de Servicio y Código de Honor

Política de privacidad

Políticas de Accesibilidad

Política de marcas

Mapa del Sitio

Cookie Policy

**Your Privacy Choices** 

# **Contáctanos**

Centro de ideas

Contáctenos

Centro de Ayuda

Security

Kit Multimedia













© 2025 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 粤ICP备17044299号-2