**TP#3: EDA Indicadores de Calidad de Vida Mundial**

Bienvenidos a nuestro tercer TP!! Un increíble DataSet con muchísimos datos sobre diferentes indicadores de calidad con información para los diferentes países del mundo…

**Material a trabajar para nuestra próximo encuentro**

El DataSet completo pueden descargarlo de este link:  
<https://mega.nz/file/gx00xSpb#IwxEcT72T-eUhMGZARN7lyiG-zOcwBECE3rz3Mw_YT8>

Como verán, tienen información distribuida en diferentes archivos **.CSV**: Sólo uno contiene los datos a analizar (**HNP\_StatsData.csv**) y el resto contienen información que nos permite entender lo que leemos allí... Hay referencias sobre los indicadores tabulados, su significado y su manera de cálculo; referencias sobre los países y qué datos se disponen para cada país, etc... Deberán mirar cada tabla para entender su estructura y la información que nos aporta.

Puntualmente, para este DataSet, tengo la intención que puedan poner en práctica las herramientas específicas de **pandas** para poder manipular datos en columnas, ya que habitualmente los habríamos encontrado en filas con la clave: {"Country Code","Indicator Code","Year"}

**Un Tip para tener presente**: vean el método **.melt** de pandas... les permitirá acomodar los datos, trasponiendo "sólo las columnas necesarias"!! Pueden dejar la clave que identifica al país e indicador que nos interesa, y agregar una columna para el año y otra para el valor registrado en dicha posición!

Tener en cuenta la siguiente pauta para el trabajo:

**1.** Considerar los siguientes índices \*\* :

* SP.DYN.CBRT.IN
* SP.DYN.LE00.IN
* SP.HCI.OVRL
* SP.GNP.PCAP.CD

*\*\* Averiguar a qué se refiere cada uno de ellos...*

**2.** Quedarnos sólo con las filas del DataSet principal que contienen los datos indicados.

**3.** Establecer una “comparativa gráfica” de Argentina respecto de otros países para estos indicadores… Habrá que bucear un poco en los datos para ver qué y cómo pueden comparar. También deberemos elegir contra qué países comparar en función de lo que nos interese observar: deberá cada uno justificar su elección.

**4.** Procurar establecer una suerte de correlación “visual” entre ellos. Por ejemplo, me parece que puede ser interesante revisar los primeros dos respecto del tercero (Tasa de Natalidad y Mortalidad vs Índice de Capital Humano); o también entre el tercero y el cuarto (Índice de Capital Humano vs. Ingreso Bruto per Cápita).

Finalmente, a lo anterior agregar un último desafío: puede ser de difícil lectura una gráfica que incluya tantas columnas de datos… estamos hablando de casi 60! Una lectura visualmente más abarcable puede lograrse si recortamos en periodos más grandes, como por ejemplo, por lustros (agrupando o tomando información cada 5 años). Ya tienen las herramientas para generar esto también!

Deseándoles una linda semana, los saludo con un abrazo!!