En este primer ejercicio vas a entender la tabla de datos que estamos manejando.

Estos son los pasos que seguiremos durante esta hoja de trabajo:

1. Abre la tabla de datos con el software estadístico SPSS o más fácil la abres con Excel
2. Entiende y clasifica las variables de la tabla de datos en:
   1. Cuantitativas: discretas o continuas
   2. Cualitativas: nominales u ordinales
3. Empieza a jugar con la tabla de datos e identifica 3 aspectos que quieres ver / analizar según las variables que tenemos

¿Preparado para empezar?

*Aviso: el ejercicio lo vamos a hacer en Excel y SPSS. Pero puedes utilizar RStudio, RCommander o cualquier otro software. No hay ningún problema.*

*El ejercicio trata de entender conceptos estadísticos de forma práctica y entender un método de trabajo ☺*

*Puedes conseguir una copia de prueba SPSS durante un mes en esta página:*

[*https://www.ibm.com/analytics/es/es/technology/spss/spss-trials.html*](https://www.ibm.com/analytics/es/es/technology/spss/spss-trials.html)

*O utilizar PSPP que es el software libre y gratuito casi igual que SPSS:*

[*https://www.gnu.org/software/pspp/*](https://www.gnu.org/software/pspp/)

*Excel supongo que lo tendrás instalado ☺*

¡A por ello!

# Importar los datos: abre la tabla de datos “Datos.xlsx”

Abre la tabla de datos en Excel o en SPSS.

Lee atentamente la descripción de la tabla de datos en la segunda hoja “Info Variables”. Esta es la información resumida de las variables que dispones en la tabla de datos:

* **País 🡪** El nombre del país que estamos estudiando - la observación
* **Región 🡪** Región del mundo: Africa, Asia, Carribean, Europe, Latin Amer, North America, NorthAtlantic, Oceania
* **group** 🡪 si pertenecen al OECD (organización para el desarrollo económico y cooperativo) - oecd = pertenece, africa = no pertenece y es un país africano, other = NO pertenece y es NO africano
* **fertility** 🡪 ratio de fertilidad del país. Número de niños por mujer
* **ppgdp** 🡪PIB per cápita en Dollars USA
* **lifeExpF** 🡪Esperanza de vida media en el periodo del estudio
* **pctUrban** 🡪 porcentaje de la población en zona urbanizadas
* **infantMortality** 🡪 ratio de mortalidad infantil - muerte de un niño menor que 1 año / 1000 nacimientos

pais : chr [1:193] "Afghanistan" "Albania" "Algeria" "Angola" ...

$ region : chr [1:193] "Asia" "Europe" "Africa" "Africa" ...

$ group : chr [1:193] "other" "other" "africa" "africa" ...

$ fertility : num [1:193] 5.97 1.52 2.14 5.13 2.17 ...

$ ppgdp : num [1:193] 499 3677 4473 4322 9162 ...

$ lifeExpF : num [1:193] 49.5 80.4 75 53.2 79.9 ...

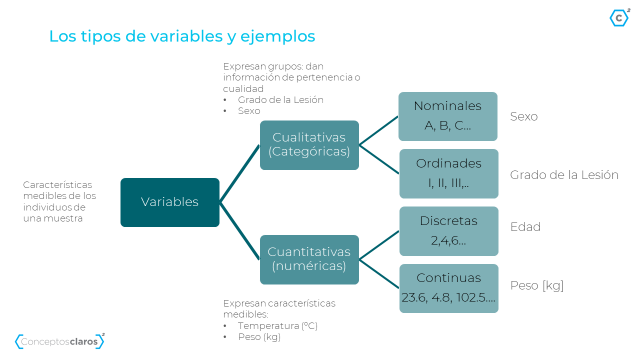
$ pctUrban : num [1:193] 23 53 67 59 93 64 47 89 68 52 ...

$ infantMortality: num [1:193] 124.5 16.6 21.5 96.2 12.3 ...

# Clasifica las variables

|  |  |  |  |
| --- | --- | --- | --- |
| Nombre Variable | ¿Qué característica mide? | ¿Qué unidades tiene? | ¿Qué tipo de variable es? |
| pais | Ninugna es el nombre del pais | NA | CUALITATIVA – nomibal NOMBRE DEL PAIS |
| region | NA | NA | CATEGORICA – AGRUPA AL CONTINENTE |
| group | NA | NA | CATEGORICA E INDICA SI PERTENCE A OCDE, AFICA O A NINGUNO |
| fertility | EN PROMEDIO CUANTOS NIÑOS DA A LUZ UNA MUJER | NIÑOS | CONTINUA |
| ppgdp | PIB PER CAPITA EN DOLLARES | MONEDA | DISCRETA |
| lifeExpF | Esperanza de vida en años | años | Continua |
| pctUrban | % de la población que vive en zonas urbanas | NA | Continua |
| infantMortality | Razón de mortalidad infantil < 1 año por cada 1000 | Niños fallecidos | continua |

Para poder rellenar la tabla inspírate en lo siguiente:



Recuerda:

* 1. Cuantitativas: discretas o continuas
  2. Cualitativas: nominales u ordinales

# Plantea el objetivo del estudio / proyecto

¡Bien!

Ahora que ya has empezado a entender la tabla de datos es momento de pensar un poquito sobre el trabajo que vamos a realizar. Es decir, qué queremos ver exactamente.

* ¿Qué podemos analizar con esta tabla de datos?
  + **Comparar pib por continentes y por si pertenecen al ocde o no.**
  + **Comparar esperanza de vid por continentes**
  + **Buscar si hay asociación entre pib y esperanza de vida**
  + **Buscar asociación entre % zonas urbanizadas vs esperanza de vida**
  + **Comparar la fertilidad entre continentes y si pertenencen al ocde.**
  + **otras**

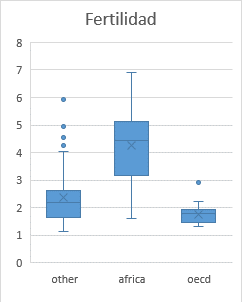
* ¿Qué objetivo perseguimos con esta tabla?
  + **Claramente parece que se busca estudiar u observar la calidad de vida en diferentes países y su asociación por continente y a que grupo pertenecen**

Para ayudarte te voy a recomendar pequeños ejercicios de visualización de datos. Según algunos objetivos que pueden ser interesantes:

## Objetivo 1 - comparando las características según si es un país del OECD:

Se trata de analizar las diferencias entre los grupos de OECD. Crea un boxplot por grupos como este. Si no sabes que es un boxplot aquí tienes un poco de información:

<https://conceptosclaros.com/como-ser-mas-rapido-boxplot-estadistica-descriptiva-parte2/>



Utiliza los vídeos Tech Zen que tienes en la página del ejercicio. Te recomiendo usar SPSS para crear estos boxplots. Es más fácil ☺

Crea los demás boxplots para poder visualizar las otras medidas de ppgdp, lifeExpF, pctUrban, infantMortality en función del grupo

|  |
| --- |
| Copia los boxplot de las demás variables |

¿Qué conclusiones obtienes de los gráficos que has creado?

Si no entiendes los gráficos no te preocupes que te desvelaré el ejercicio resuelto.

|  |
| --- |
| Claramente se puede ver que los países agrupados en la OCDE presentan mejores condiciones de vida en todos los aspectos( pib, personas en zonas urbanas, expectativas de vida,menor mortalidad infantil) , seguidos por otros y luego Afica.  Africa tiene un mejor control de la Natalidad ya que a pesar de que no tienen la condiciones de vida adecuadas resulta que las mujeres tienen mas hijos e igualmente mucha mas mortalidad de niños |

## Objetivo 2 – ¿qué variables están relacionadas con la muerte infantil?

Podemos también analizar qué variables influyen en la mortalidad infantil o qué variables están relacionadas con esta.

Crea un diagrama de dispersión o scatterplot

entre la mortalidad infantil y la esperanza de vida.

Aquí un ejemplo:

|  |
| --- |
| Copia los diagramas de dispersión de la mortalidad infantil en función también del resto de variables: ppgdp, pctUrban, fertilidad            Aquí visualizo la variable Infan Mortality con LN del PIB pero total, es decir sin desgrupar el scatter plot |

¿Qué conclusiones obtienes de los gráficos que has creado?

Si no entiendes los gráficos no te preocupes que te desvelaré el ejercicio resuelto.

|  |
| --- |
| En todos los casos parece exitir una asociación entgre las variables muertes infantiles vs:   * fertilidad (Asociacion Muy Alta y Positiva),   + "r de InfantMortality vs Fertilidad 0.857 * Poblacion Urbana (Asoc Baja. Negativa u Opuesta)   + "r de InfantMortality vs Poblacion Urbana -0.599 * PIB (Asoc.Baja Negativa u Opuesta)   + "r de InfantMortality vs PIB -0.51 * Esperanza de Vida (Asoc. Negativa u Opuesta   + "r de InfantMortality vs Esperzana de Vida -0.93 |
|  |

## ¿Qué otros aspectos podemos analizar de esta tabla de datos?

Lista otros aspectos que te gustaría encontrar y analizar de esta tabla de datos.

|  |
| --- |
| Ver esta misma asociación pero por el grupo al que pertenecen los países. Sobre todo porque no me queda clara la asociación en la gráfica de Muertes infantiles vs Población Urbana y PIB  Para el caso de Muertes Infantiles vs PIB al ser africa un grupo con bajo PIB no parece que exitiese asociación. Es baja, Sin embargo en el caso de otros se nota que a mayor pib menos muertes infantiles, pareciera estar un poco mas asociadas. |

¡Enhorabuena por el trabajo!

Es una primera toma de contacto para ir jugando con la tabla de datos.

Haciendo estos pequeños ejercicios de ayudará a avanzar

¡Ciao, ciao!