

MEMORIA EDA GONZALO PÉREZ DIÉZ

Descripción:

El trabajo está hecho con datos recogidos del gobierno, del ine y de páginas de confianza.

En él se trata de visualizar la distribución de los recursos sanitarios actuales a lo largo del territorio.

Limpieza de datos:

En un principio tenía datos sobre el número de muertes según el tipo de muerte, sin embargo, los datos eran de tan baja calidad que busque otros datos con los que trabajar. Llegué a cambiar tres veces de idea sobre el proyecto hasta encontrar los datos actuales.

Los datos que recogí descargando los csv estaban limpios, aunque modifiqué la columna de las comunidades autónomas, cambié el nombre de las columnas para hacerlas parejas a las columnas de la base de datos donde más tarde introduje los datos y recogí manualmente algunos de los datasets (poblacion, poblacion_mujeres, PIB y gasto en sanidad).

Los datos los pasé a una base de datos en SQLite debido a la facilidad con la que se pueden hacer consultas y a que, al realizar dichas consultas no modificabas el dataframe original. También elegí esta opción porque estaba bastante familiarizado con SQL.

Para saber más sobre la limpieza de datos y ver todo el proceso ir al apartado 'notebooks' del zip.

Enlaces:

PIB por CCAA: <https://datosmacro.expansion.com/pib/espana-comunidades-autonomas>

Población total y de mujeres por CCAA: <https://www.ine.es/jaxiT3/Tabla.htm?t=2853&L=0>

Gasto sanidad <https://www.mscbs.gob.es/estadEstudios/estadisticas/docs/EGSP2008/egspPrincipalesResultados.pdf>

datos_hospitales: <https://www.mscbs.gob.es/ciudadanos/centrosCA.do>

url_gasto_sanidad: <https://www.mscbs.gob.es/estadEstudios/estadisticas/inforRecopilaciones/docs/presupuestosIniciales.pdf>

url_población: https://es.wikipedia.org/wiki/Anexo:Comunidades_y_ciudades_aut%C3%B3nomas_de_Espa%C3%B1a