

Instrucciones:

- Folio. Cada folio tiene varias viviendas, pero siempre tiene un hogar. Entonces, podríamos decir que folio, número de vivienda y número de hogar son únicos, excepto cuando se produce una repetición, donde la salvedad queda dada por la persona.
- La variable persona solo existe en el formulario población. Si voy al formulario de vivienda y hogar no va a existir, por lo tanto, ¿dónde debería tratar de hacer un join pensando en cuál es la base más desagregada? La base que tiene más registros es POB (17 millones de registros que corresponden a cada una de las personas censadas).
- Sobre eso, van a haber personas por hogar y dentro de la vivienda habrá hogares (ejemplo: en una vivienda puede vivir más de una familia).
- Tenemos el folio que se repite muchas veces y después tenemos viviendas, dentro de ellas hogares y personas. Si abro otra base veré que el NPER no existe, por lo tanto, ¿qué me permitiría juntar casos? Las tres variables que se repiten: folio, nviv y nhogp.
- La complejidad de la tarea es que si yo quiero crear una nueva variable que me permita tener “este” identificador único tengo que ver cuál es el recorrido de la variable: el número de folio es fijo, siempre tiene la misma cantidad de caracteres. En cambio, si me pongo en la vivienda y la pongo en forma descendente encontraré un folio con 907 viviendas, y, por lo tanto, tengo que dejar espacio para generar esa variable. El truco es identificar cuál es el número más grande de vivienda y cuál es el número más grande de hogar. Así se puede crear un gran y hermoso código único.
- Fórmula de manzana:
$$MZ+ENTIDAD + 1.000*(ZONA+LOC) + 1.000.000*(AREA) + 10.000.000*DTO + 1.000.000.000*COM$$

En SPSS no existe ni la región ni la provincia, pero sí existe la comuna, entonces ésta se transforma en la unidad máxima con la que trabaja a nivel político administrativo. La fórmula termina con la comuna multiplicada por un número gigante, ese número es el espacio que le dejo después de la comuna al código para después sumarle todo el resto.
- Entonces, después de eso, debo hacer lo siguiente:
Datos → fusionar archivos.
- La idea para la próxima clase es intentar traer el censo 2012 fusionado que tenga la población, la vivienda, los hogares y el código de manzana creado.