



Visualización de datos censales en R

Juan Fernando Correa Caicedo
correaifc@gmail.com
Foromundo UNIGIS
Junio 2017

Contenido

- Objetivos
- Flujo de trabajo
- Herramientas
- SIG vs R
- Representación de datos espaciales
- El problema (reto)
- Cartografía y datos censales
- Data Frame
- SP
- tidyverse
- ggplot2
- Bibliografía



Objetivos

- Identificar fuentes de datos y cartografía censal
- Introducir conceptos básicos y estructuras de datos para el manejo de datos espaciales y alfanuméricos en R
- Elaborar gráficas y mapas usando ggplot2
- Manipular y calcular datos con R
- Compartir experiencias y aprendizajes sobre el uso de datos del Censo de Población 2005 y su visualización.
- Usar los datos del censo 2005

El flujo de trabajo

¿Qué? ¿Para qué? ¿Para quién?
¿tiempo ¿presupuesto?

Disponibilidad, resolución,
calidad, actualización,
documentación

preparar, depurar, transformar,
explorar, calcular, analizar

¿usuarios? ¿formatos?
¿impreso digital? ¿interactividad?
Diseño

Requerimientos

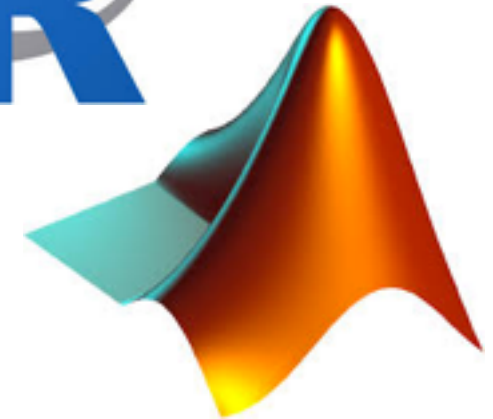
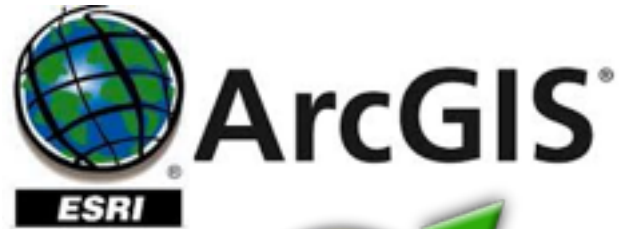
Recopilar
datos

Procesamiento

Visualización



Herramientas



SIG versus R

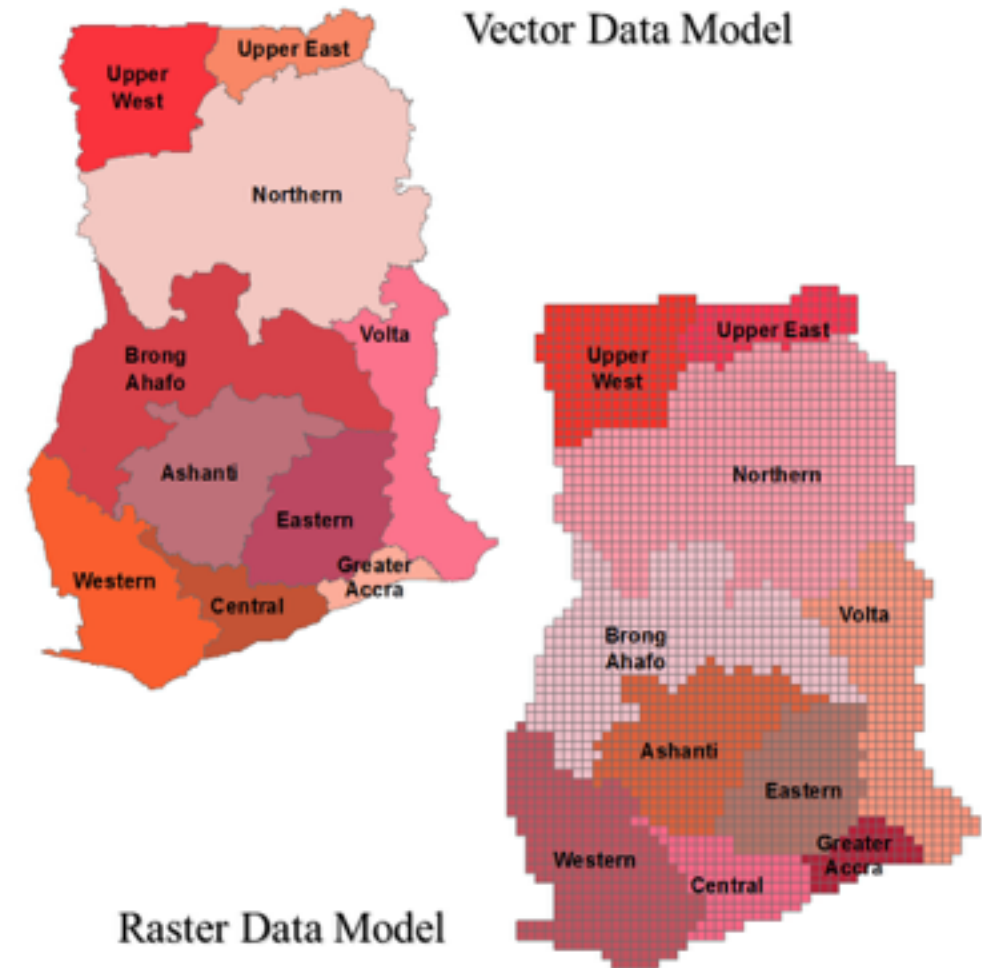
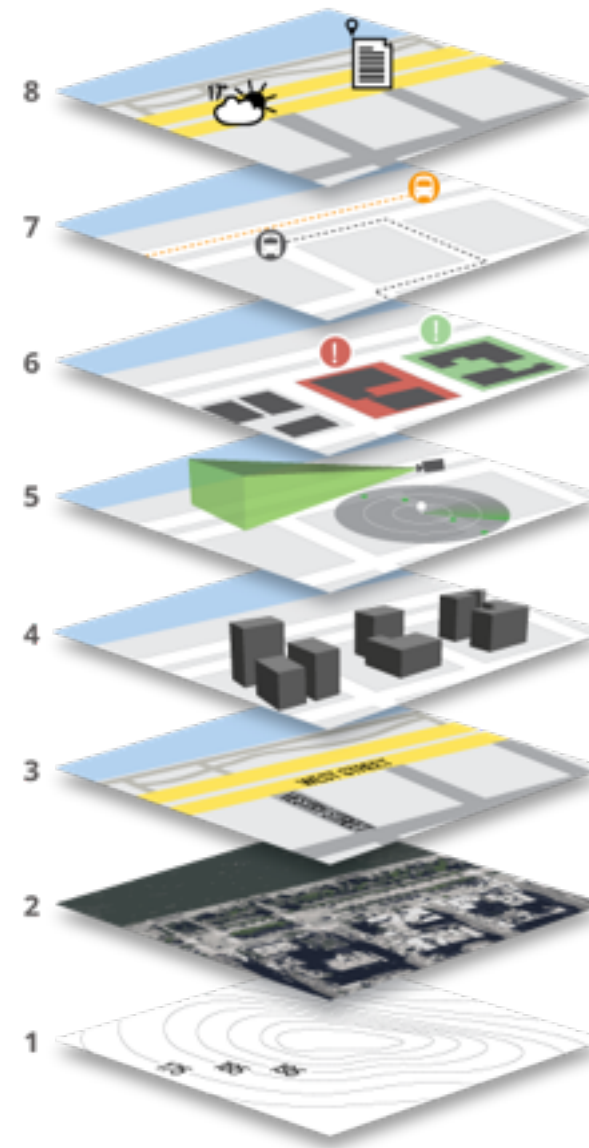
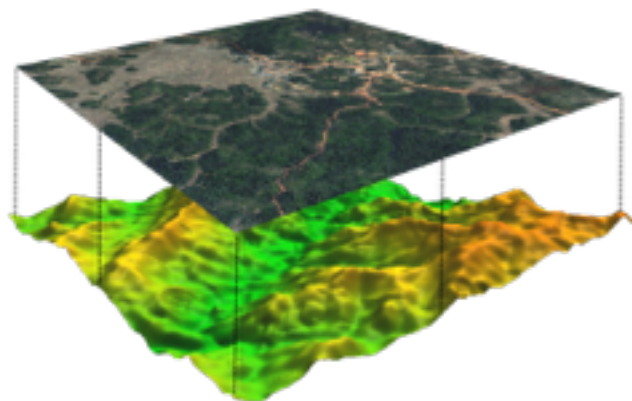
- Interfaces de usuario y herramientas parametrizadas
- Interacción visual
- Gestión de datos (bases de datos espaciales)
- Operaciones geométricas
- Flujos de trabajo estándar
- Producción de mapa único
- Haga clic en, click, click,...
- Énfasis cartográfico

- Interfaz es el texto, lenguaje de programación.
- Enfocado a los datos y los modelos
- Análisis matemático y procesamiento flexible
- Creatividad e innovación
- Muchos mapas (más simples)
- Repetibilidad (script)
- Software libre (gratis)
- Integración con otros análisis
- Desarrollo de nuevos métodos



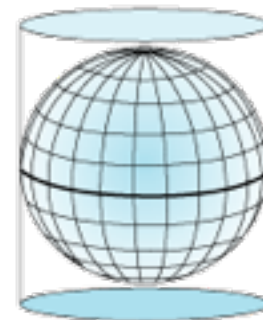
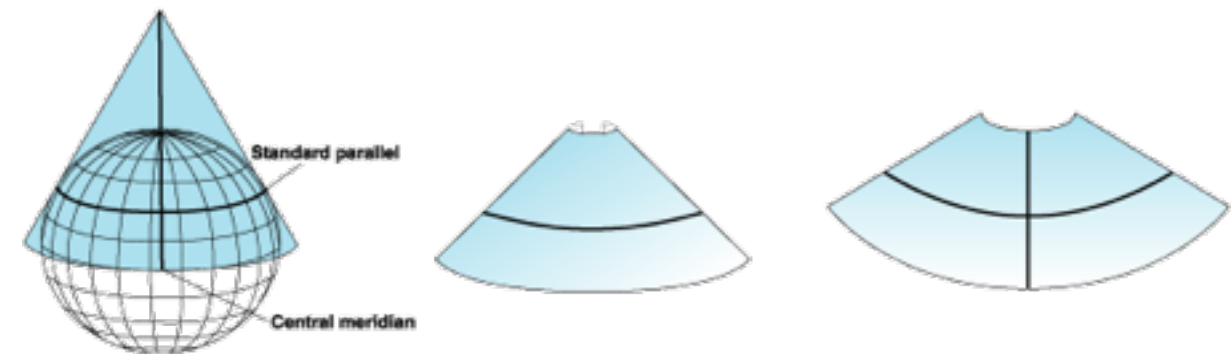
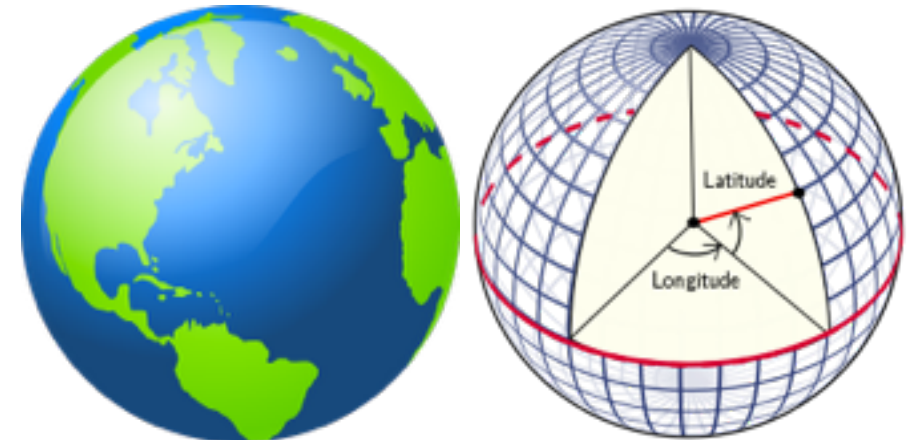
Representación de datos espaciales

- Representaciones vectoriales (features con geometría y límites definidos)
- Datos tipo raster (variables continuas sobre la superficie)

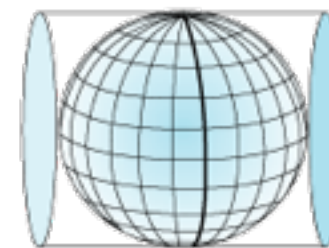


Coordenadas geográficas y proyecciones cartográficas

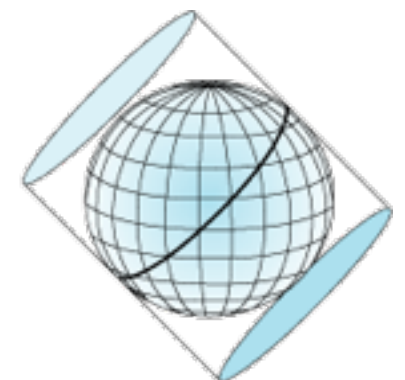
- **Coordenadas geográficas:** coordenadas esféricas, se usa para definir puntos sobre una superficie esférica especial, la terrestre.
(“World Geodetic System”, 2017)
- **Proyección cartográfica:** sistema de representación gráfica que establece una relación ordenada entre los puntos de la superficie curva de la Tierra y los de una superficie plana (mapa). Estos puntos se localizan auxiliándose en una red de meridianos y paralelos, en forma de malla.
(“Proyección cartográfica”, 2017)



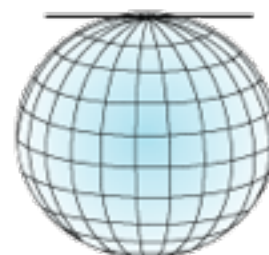
Normal



Transverse



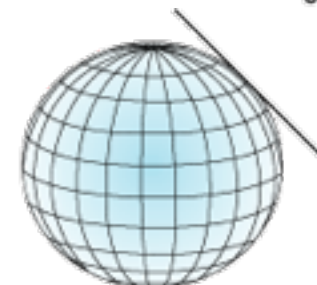
Oblique



Polar



Equatorial



Oblique

WGS84 y el Magna-Sirgas Cali

El WGS84 es un sistema de coordenadas geográficas mundial que permite localizar cualquier punto de la Tierra (sin necesitar otro de referencia) por medio de tres unidades dadas. WGS84 son las siglas en inglés de World Geodetic System 84 (que significa Sistema Geodésico Mundial 1984).
("WGS84", 2017)

Sistema de Coordenadas Cartesiano de Cali

- Proyección: Transversa Mercator.
- Elipsoide: GRS 1980.
- Datum: MAGNA Origen Santiago de Cali.
- Latitud: 3°26'30,779" Norte.
- Longitud: 76°31'14,025" Oeste.
- Falso Norte: 872.364,630 m.
- Falso Este: 1'061.900,180 m.
- Plano de Proyección: 1000 m.

("Proyecciones y Transformaciones Cartográficas", s/f)

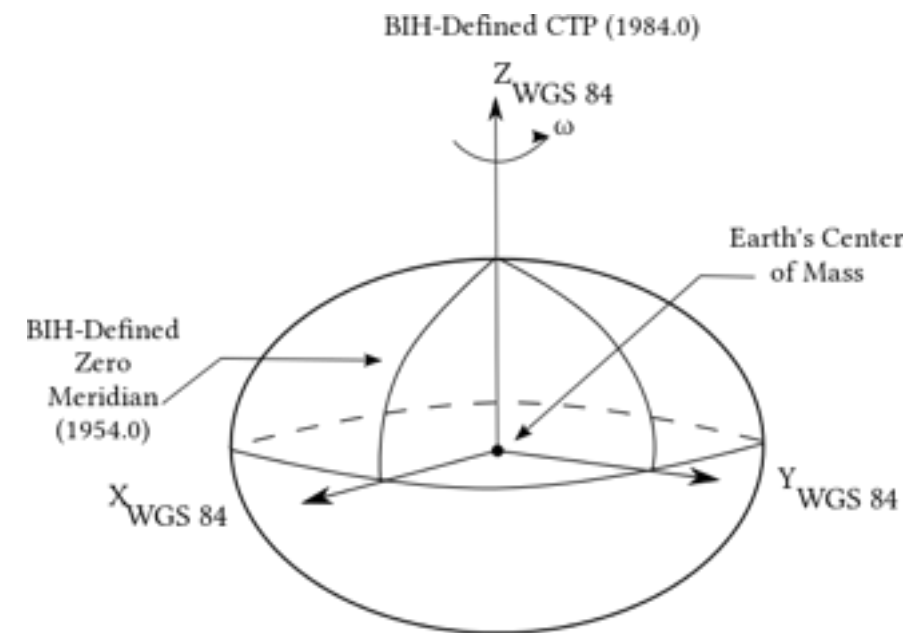
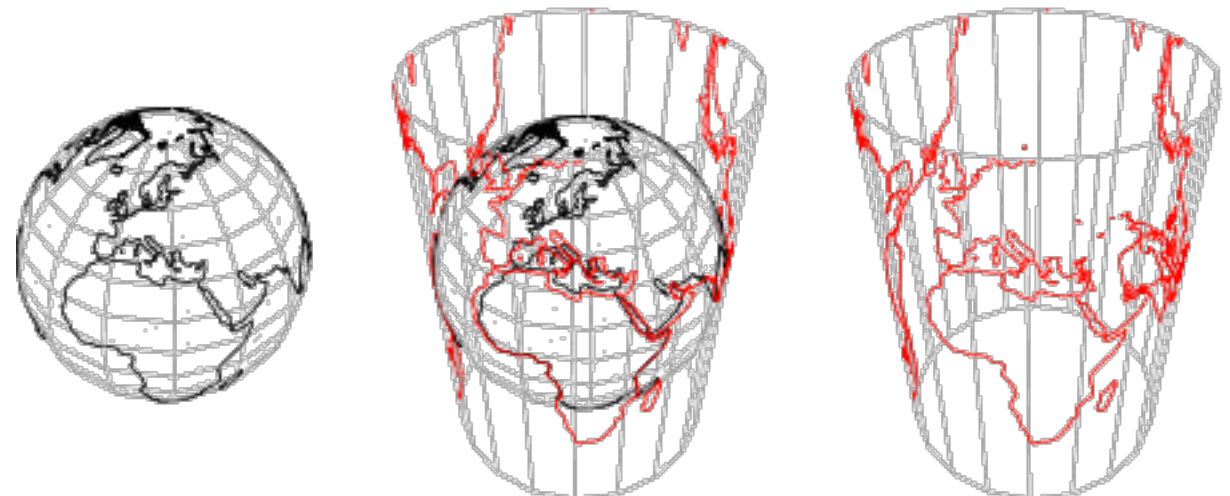


Figure 1.1 WGS 84 Reference Frame



- Generar mapas y gráficas estadísticas que permitan caracterizar la densidad de población de la ciudad de Santiago de Cali y los tipos de viviendas, usando sectores censales
- Formato digital para incluir en informes y distribuir en la web.

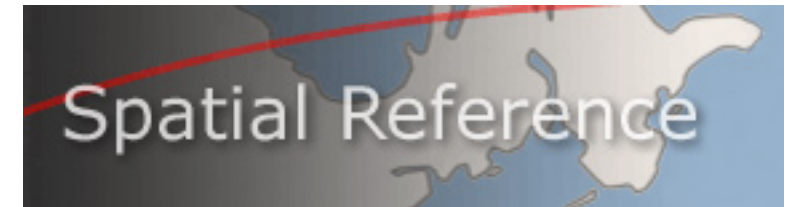
Cartografía y datos censales

Recopilar datos

- Fuentes



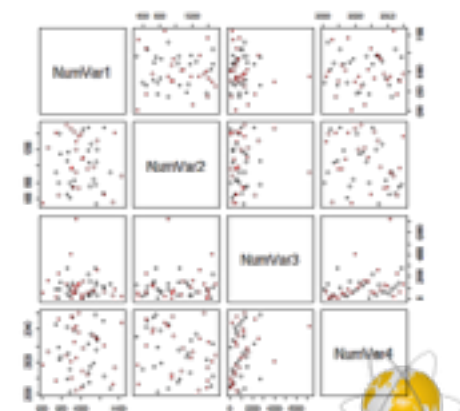
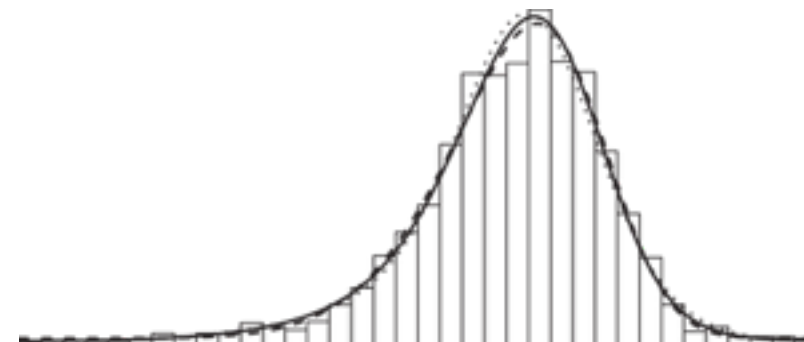
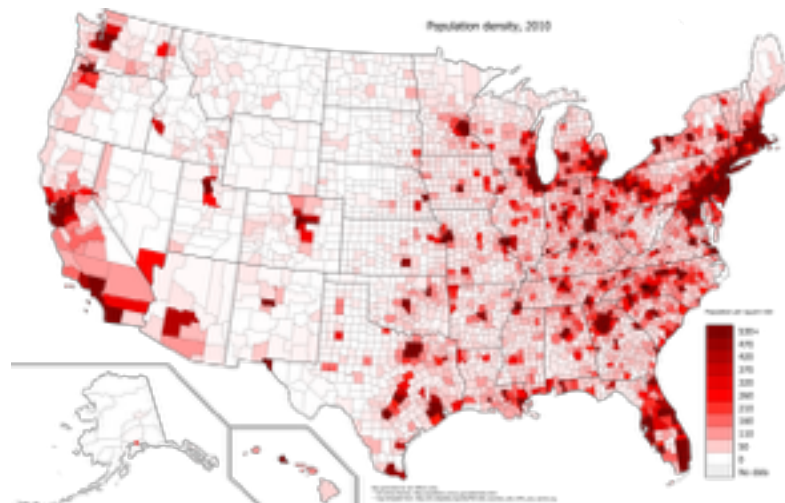
Geoportal



- Documentación

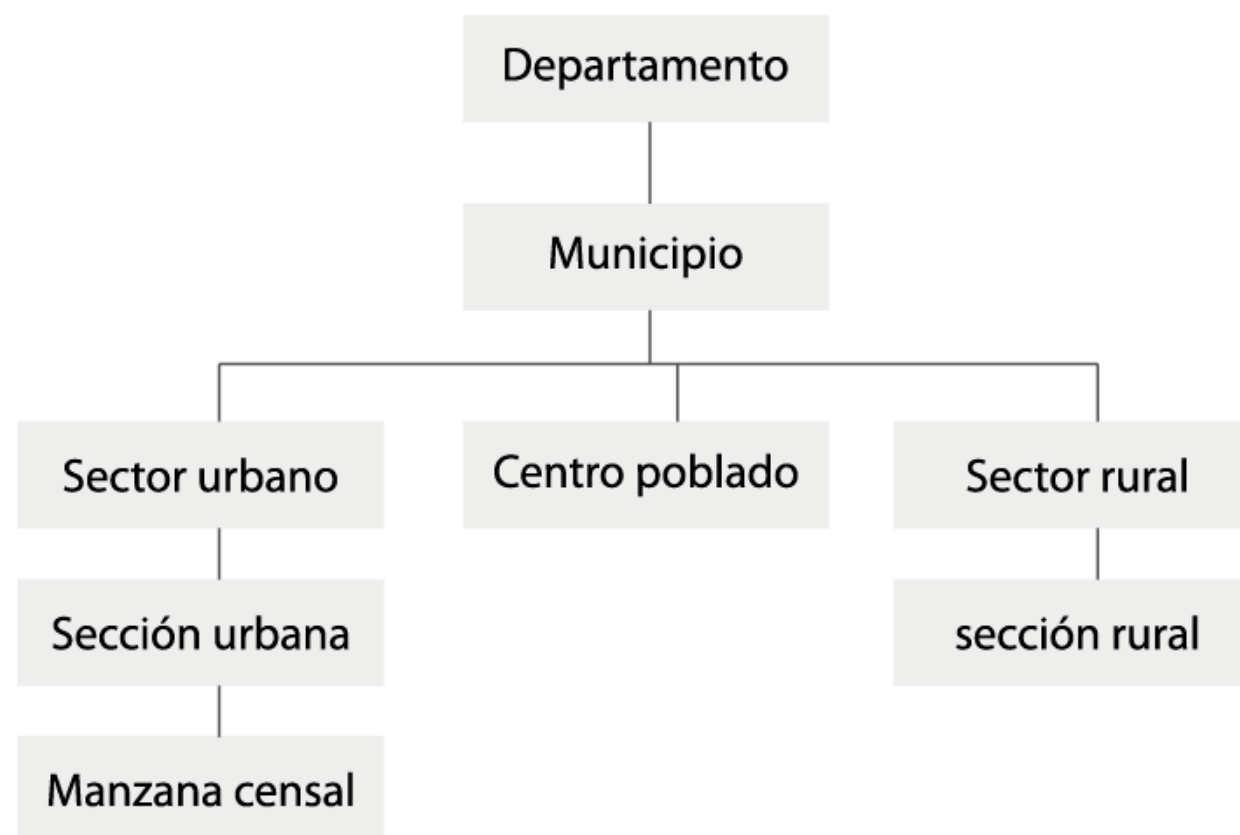


- Datos



El censo de población y la cartografía MGN

Marco Geoestadístico Nacional: Sistema diseñado por el DANE, para referenciar la información estadística con los lugares geográficos correspondientes, con el propósito de facilitar los procesos de planeación, recolección de datos, procesamiento, análisis y difusión de la información.



Niveles Urbanos del MGN

Unidad Geo	Definición
Centro Poblado	Es un área con características urbanas, ubicada en el espacio rural de un municipio, conformado por 20 o más viviendas contiguas o adosadas entre sí.
Sector Urbano	División cartográfica censal a nivel urbano, equivale por lo general a un barrio (en las ciudades principales). Comprende entre 1 y 9 secciones urbanas.
Sección Urbana	División cartográfica urbana delimitada con fines operativos censales, que equivale aproximadamente a 20 manzanas censales.
Manzana Censal	Lote de terreno edificado o sin edificar delimitado por vías de tránsito vehicular o peatonal de carácter público.

- Para el caso de descarga de información con agregaciones sector urbano, sección urbana o manzana censal, es importante suprimir los dígitos 7 y 8 del código censal antes de cruzar con la información de las capas correspondiente de la cartografía censal del MGN

tributos de los Sectores Urbanos

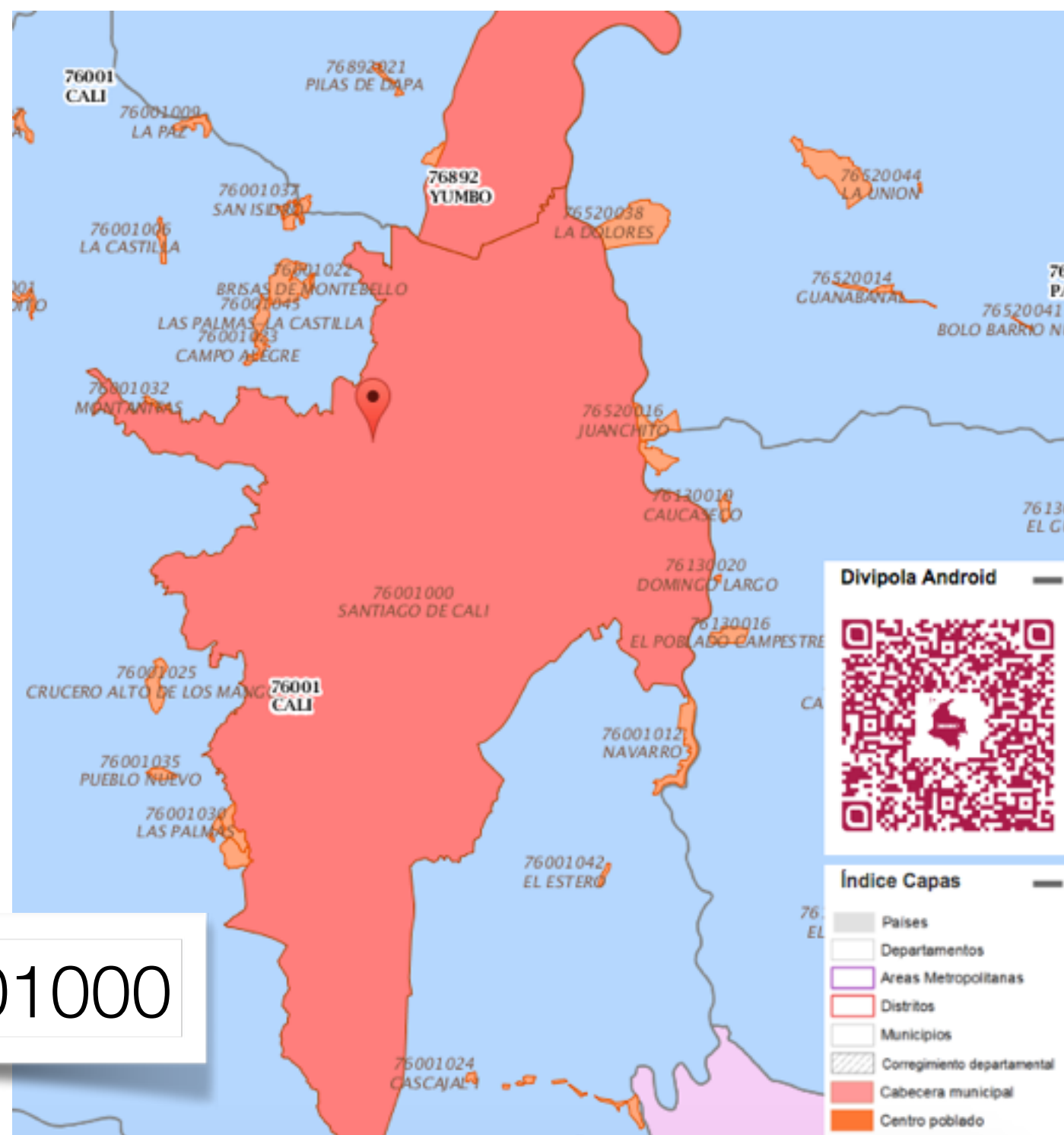
Nombre atributo	Estructura del campo		Opcional	Descripción del atributo
SETR_CLSE_	Texto	2	NO	Código del departamento
SECR_SETR_	Texto	3	NO	Código que identifica al municipio
CPOB_SECR_	Texto	1	NO	Tipo de centro poblado
CPOB_SECR1	Texto	3	NO	Código que identifica el sector rural
CPOB_SEC_1	Texto	2	NO	Código con el que se identifica la sección rural
CPOB_CPOB_	Texto	3	NO	Código que identifica el centro poblado
SETU_CCDGO	Texto	4	NO	Código del sector urbano
SETU_CCNCT	Texto	18	NO	Código DANE concatenado a nivel sector urbano
SETU_NAREA	Numérico	12,3	NO	Área en metros cuadrados oficial del sector urbano
SETU_CSMBLO	Texto	3	NO	Símbolo con el que se identifica el sector urbano en las salidas gráficas
SETU_NANO	Entero	4	NO	Año de la vigencia de la información

Divipola


La División Político-administrativa de Colombia Divipola es un estándar de codificación que permite contar con un listado organizado y actualizado de la totalidad de unidades en que está dividido el territorio nacional, dándole a cada departamento, municipio, corregimiento departamental y centro poblado, el máximo de estabilidad en su identificación.

Cali municipio: 76001

Cali Centro poblado: 76001000



Sistema de consulta del Censo 2006



Sistema de Consulta

CENSO

Censo Básico

Definición de Parámetros

Título de la Tabla

Seleccione una Variable:

Sexo

Edad en grupos quinquenales

Edad en grupos decenales

Edad para Personas de 18 Años y más

Parentesco con el Jefe

Donde nació

Departamento de nacimiento

Departamento del último cambio

Departamento donde vivía hace 5 años

Departamento residía la mamá al nacer

Nivel Geográfico de Salida:

Selección Geográfica:

Filtro:

Formato de Salida:

ya tenemos los datos.

Data Frame de R

- Es una estructura de datos que almacena una tabla con variables (columnas) y observaciones (filas). Cada variable puede ser de un tipo de dato distinto.

		Variables (Columnas)			
		↓	↓	↓	↓
		Pais	Censo	Población urbana	Población total
Observaciones (Filas)	→	Argentina	1980	23.198.068	27.949.480
	→	Argentina	1991	28.832.127	32.615.528
	→	Argentina	2001	32.380.296	36.260.130
	→	Argentina	2010	36.907.728	40.117.096

Valores
(celdas)

SP package

- sp: paquete que implementa métodos y estructuras de datos para el manejo de datos espaciales

matrix or data frame
with Lat Lon



`SpatialPoints`



`SpatialPointsDataframe`

matrix or data frame
with Lat Lon

`Line`



`Lines`



`SpatialLines`



`SpatialLinesDataframe`

matrix or data frame
with Lat Lon

`Polygon`



`Polygons`



`SpatialPolygons`



`SpatialPolygonsDataframe`

Estructura básica de Spatial*Dataframe

objeto espacial
(clase)

geometria

data (frame)

proyección

caja limite

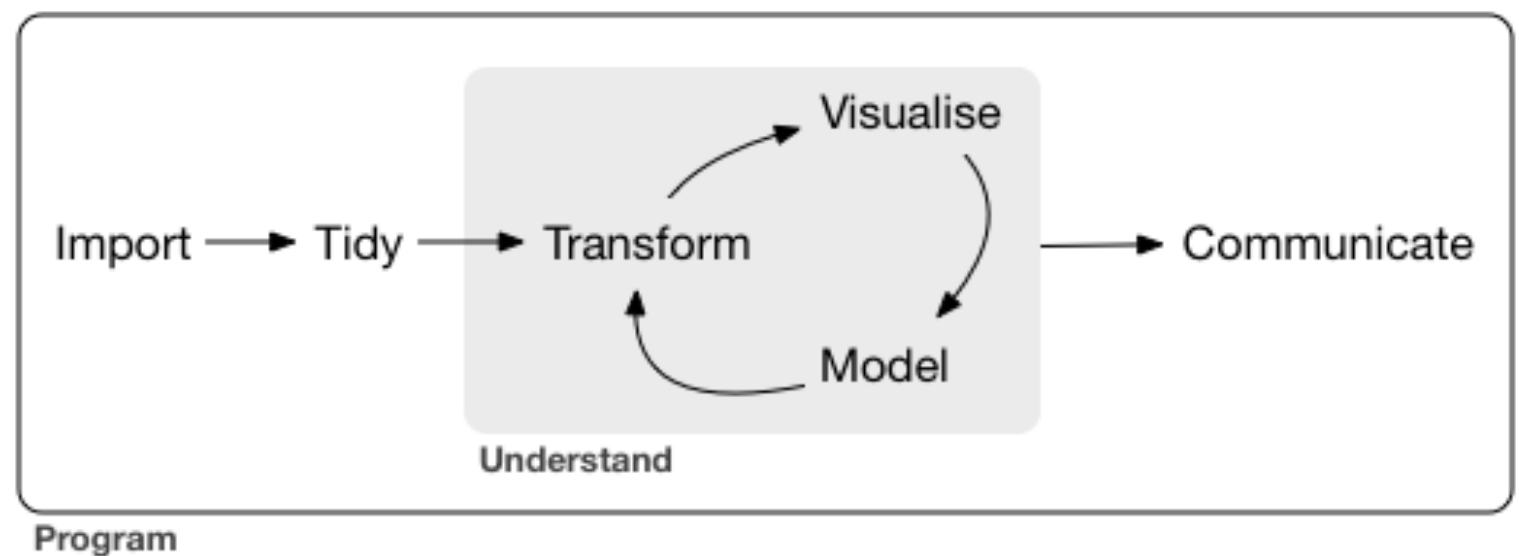
tidyverse package



Conjunto de herramientas para transformar, visualizar, modelar y comunicar datos en R.

Verbos importantes:

- select
- mutate
- rename
- summarise
- arrange
- join
- gather



ggplot2



plot = datos + Aesthetics + geometría.



manos a la obra

Recursos

- – [R] Data Visualization with ggplot2 (1)Finterest. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://www.hihaho.synology.me/hoon/?p=164>
- 6. Coordinate Reference Systems — R Spatial. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://rspatial.org/spatial/rst/6-crs.html>
- Applied Spatial Data Science with R. (2016, enero 26). Recuperado el 22 de junio de 2017, a partir de <https://blog.dominodatalab.com/applied-spatial-data-science-with-r/>
- Censo general 2005. (s/f). Recuperado el 20 de junio de 2017, a partir de <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-general-2005-1>
- Converting data between wide and long format. (s/f). Recuperado el 22 de junio de 2017, a partir de http://www.cookbook-r.com/Manipulating_data/Converting_data_between_wide_and_long_format/
- Geoportal DANE - Descargar cartografía MGN. (s/f). Recuperado el 21 de junio de 2017, a partir de <https://geoportal.dane.gov.co/v2/index.php>
- Geoportal DANE - Divipola. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://geoportal.dane.gov.co/v2/index.php>
- Geoportal DANE - Marco Geoestadístico Nacional. (s/f). Recuperado el 20 de junio de 2017, a partir de <https://geoportal.dane.gov.co/v2/index.php>
- GGally. (s/f). Recuperado el 22 de junio de 2017, a partir de http://ggobi.github.io/ggally/#life_expectancy_model_fitting

- ggplot2: Guide to Create Beautiful Graphics in R (Data visualization) (Volume 1): Alboukadel Kassambara: 9781532916960: Amazon.com: Books. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://www.amazon.com/ggplot2-Create-Beautiful-Graphics-visualization/dp/1532916965>
- Grolemund, G., & Wickham, H. (s/f). R for Data Science. Recuperado a partir de <http://r4ds.had.co.nz/>
- Home. (s/f). Recuperado el 23 de junio de 2017, a partir de <https://www.rstudio.com/>
- Introduction to R Graphics with `ggplot2`. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html>
- Introduction to Spatial Data Types in R. (s/f). Recuperado el 22 de junio de 2017, a partir de https://cengel.github.io/rspatial/2_spDataTypes.nb.html
- Lovelace, R., & Cheshire, J. (2014). Introduction to visualising spatial data in R. Recuperado a partir de <http://eprints.ncrm.ac.uk/3295>
- MAGNA_Colombia_Cali: SR-ORG Projection -- Spatial Reference. (s/f). Recuperado el 21 de junio de 2017, a partir de <http://spatialreference.org/ref/sr-org/7393/>
- mapping - Simple way to subset SpatialPolygonsDataFrame (i.e. delete polygons) by attribute in R - Stack Overflow. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://stackoverflow.com/questions/13443372/simple-way-to-subset-spatialpolygonsdataframe-i-e-delete-polygons-by-attribute>
- Mapping in R using the ggplot2 package. (2014, julio 16). Recuperado el 22 de junio de 2017, a partir de <http://zevross.com/blog/2014/07/16/mapping-in-r-using-the-ggplot2-package/>

- Mauricio. (2016, diciembre 22). Curso de R | Manipulación de data frames con “tidyr”. Recuperado el 22 de junio de 2017, a partir de <https://mauricioanderson.com/curso-r-tidyr/>
- proj.4 — proj.4 4.9.3 documentation. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://proj4.org/>
- Proyección cartográfica. (2017, junio 20). En Wikipedia, la enciclopedia libre. Recuperado a partir de https://es.wikipedia.org/w/index.php?title=Proyecci%C3%B3n_cartogr%C3%A1fica&oldid=99972863
- Proyecciones y Transformaciones Cartográficas. (s/f). Recuperado el 21 de junio de 2017, a partir de http://www.cali.gov.co/publicaciones/105289/proyecciones_transformaciones_cartograficas_idesc/
- r - How to customize lines in ggpairs [GGally] - Stack Overflow. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://stackoverflow.com/questions/30858337/how-to-customize-lines-in-ggpairs-ggally>
- R filter shapefiles with dplyr - Geographic Information Systems Stack Exchange. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://gis.stackexchange.com/questions/156268/r-filter-shapefiles-with-dplyr>
- RPubS - Geospatial Data Processing and Analysis in R. (s/f). Recuperado el 22 de junio de 2017, a partir de <https://rpubs.com/ajlyons/rspatialdata>
- Servicios WFS - Web Feature Service. (s/f). Recuperado el 21 de junio de 2017, a partir de http://www.cali.gov.co/publicaciones/46986/servicios_wfs_idesc/
- Spatial Data in R: Vector Data. (s/f). Recuperado el 22 de junio de 2017, a partir de http://www.nickeubank.com/wp-content/uploads/2015/10/RGIS1_SpatialDataTypes_part1_vectorData.html

- tidyverse website • tidyweb. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://tidyverse.org/>
- Tipos de proyección—Ayuda | ArcGIS for Desktop. (s/f). Recuperado el 21 de junio de 2017, a partir de http://desktop.arcgis.com/es/arcmap/10.3/guide-books/map-projections/projection-types.htm#ESRI_SECTION2_B646A44462A94F5B8AB25D244B899DEA
- Urban Spatial. (s/f). Recuperado el 22 de junio de 2017, a partir de <http://urbanspatialanalysis.com/dataviz-tutorial-mapping-san-francisco-home-prices-using-r/>
- WGS84. (2017, mayo 1). En Wikipedia, la enciclopedia libre. Recuperado a partir de <https://es.wikipedia.org/w/index.php?title=WGS84&oldid=98788099>
- World Geodetic System. (2017, junio 11). En Wikipedia. Recuperado a partir de https://en.wikipedia.org/w/index.php?title=World_Geodetic_System&oldid=785092213

Los contribuyentes, por pagar el censo.



Agradecimiento

