Trabajo Final de Estadistica y Econometria Espacial

Wilson Guallasamin

2024-09-08

# Actividad Propuesta: Análisis Espacial de Amenidades de

Salud en Parroquias de Ecuador

**OBJETIVO :** Esta actividad te permitirá aplicar y reforzar conceptos clave en el análisis espacial, enfocándote en la relación entre las amenidades de salud y la población a nivel de parroquias en Ecuador. A través de esta práctica, tendrás la oportunidad de trabajar con diferentes fuentes de datos, realizar cálculos de indicadores relevantes, y reflexionar sobre la calidad y precisión de las fuentes de información geográfica. Además, te ayudará a comprender cómo los datos espaciales pueden integrarse y analizarse para obtener insights significativos que apoyen la toma de decisiones.

**INTRODUCCION:**

Para este analisis usaremos la provincia de Guayas . Guayas es una de las veinticuatro provincias que conforman la República del Ecuador, localizada en la región litoral del país, al suroeste del mismo. Su capital es la ciudad de Guayaquil. La provincia es el mayor centro comercial, económico e Industrial de Ecuador. Con sus 4’391.923 millones de habitantes, Guayas es la provincia más poblada del país y contiene el 25,93 % de la población de la república. Guayas se encuentra dividida políticamente en veinticinco cantones, de los cuales se derivan cincuenta parroquias urbanas y veintinueve parroquias rurales. Las actividades principales de la provincia son la industria, ya que se encuentra el puerto y las mayores fábricas en Guayaquil; el sector agropecuario en la vida rural, y el turismo, principalmente por sus famosas playas y festividades montubias.

# Instalar y cargar los paquetes necesarios # install.packages(c("sf", "tidyverse", "osmdata","spdep","cowplot))   
library(sf)   
library(tidyverse)   
library(osmdata)   
library(spdep)   
library(cowplot)  
library(lattice)  
library(readxl)  
setwd('C:/Users/WELCOME/Documents/Curso/Curso Geoestadistica/Examen Final')

# Parte 1: Identificación y Reflexión sobre las Amenidades de Salud

Para el analisis que vamos hacer, nos basamos en las herramientas de *“OpenStreetMaps”* usando *Boundary Box* e identificar las amenidades sobre la provincia del Guayas.

bbox<- getbb("Guayas, Ecuador")   
   
 osm\_query <- opq(bbox) %>%   
 add\_osm\_feature(key = "amenity", value = "hospital") %>%   
 osmdata\_sf()  
   
 puntos<- osm\_query$osm\_points

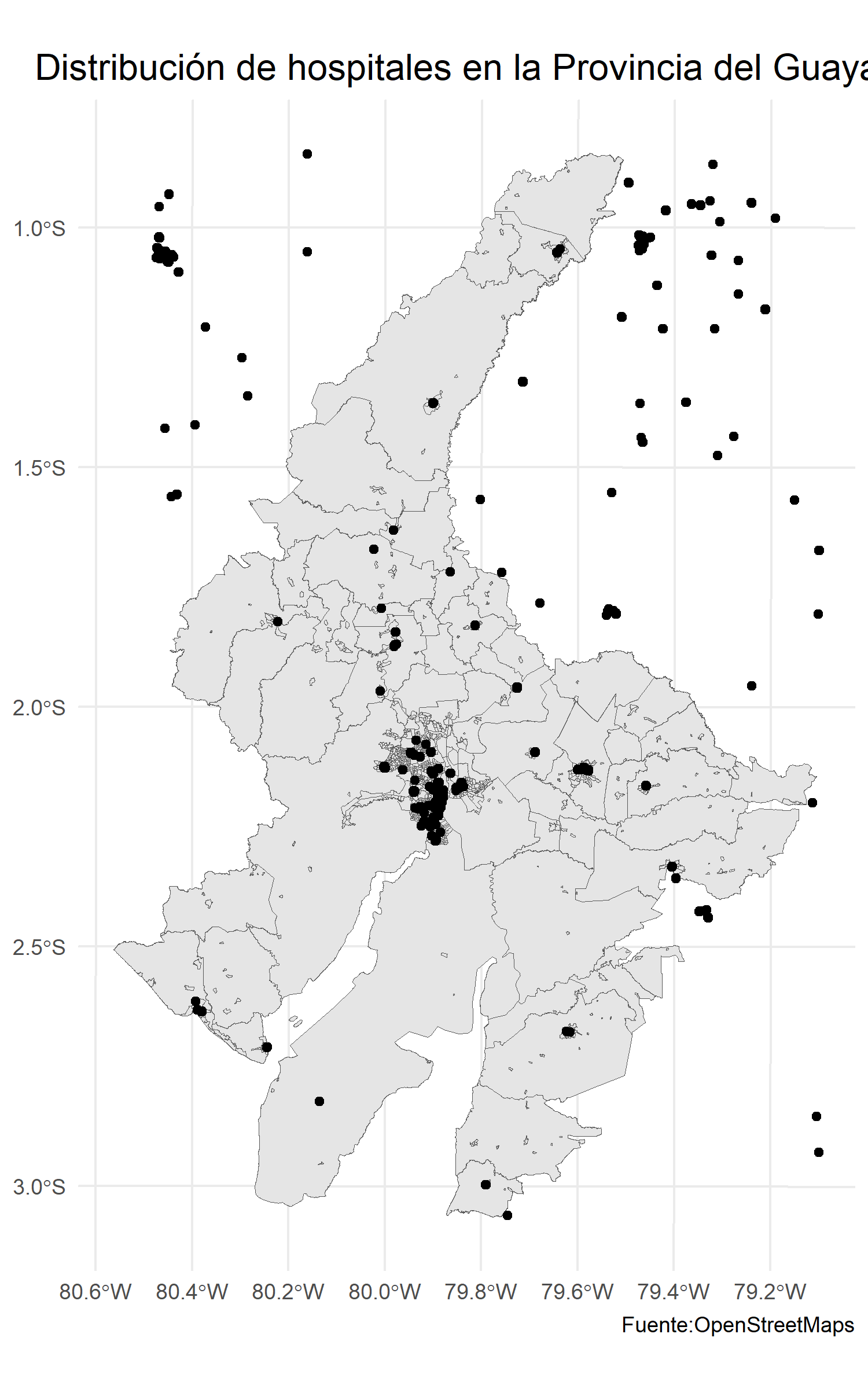
Cargamos el mapa del Ecuador y ubicamos los puntos antes encontrados sobre este mapa para tener una visualizacion general de los hospitales.

# Cargar y preparar los datos de zonas censales  
zonas\_nacional <- st\_read("C:/Users/WELCOME/Documents/Curso/Curso Geoestadistica/Examen Final/Data/GEODATABASE\_NACIONAL\_2021/GEODATABASE\_NACIONAL\_2021.gdb/", layer = "zon\_a")

Reading layer `zon\_a' from data source   
 `C:\Users\WELCOME\Documents\Curso\Curso Geoestadistica\Examen Final\Data\GEODATABASE\_NACIONAL\_2021\GEODATABASE\_NACIONAL\_2021.gdb'   
 using driver `OpenFileGDB'  
Simple feature collection with 5888 features and 6 fields  
Geometry type: MULTIPOLYGON  
Dimension: XY  
Bounding box: xmin: -685704.9 ymin: 9445216 xmax: 1147852 ymax: 10162550  
Projected CRS: WGS 84 / UTM zone 17S

zonas <- zonas\_nacional %>% filter(str\_detect(zon, "^09"))

Ubicamos sobre el mapa, los puntos de los hospitales de la provincia del Guayas para tener un vista previa de la distribución.



Vemos gran cantidad de concentración en la ciudad de Guayaquil. Aun asi, tenemos cierta concentración sobre algunas de las ciudades importantes de la provincia

# Parte 2: Unión de Polígonos de Zonas Censales a Nivel de Parroquias

Leemos la información sobre el resultado censal realizado en el año 2022.

pob\_nacional <- read\_excel("C:/Users/WELCOME/Documents/Curso/Curso Geoestadistica/Examen Final//Data/01\_2022\_CPV\_Estructura\_poblacional.xlsx", sheet = "1.2", range = "B10:G1299")  
  
names(pob\_nacional) <- c("provincia", "canton", "parroquia", "total\_nacional",  
 "hombres", "mujeres")

Como podemos notas, tenemos que realizar un fitltrado para quedarnos con la provincia del Guayas y ademas ir eliminando columnas innecesarias que forman parte del archivo resumen obtenido del INEC.

pob\_nacional <- pob\_nacional %>% filter(!is.na(provincia),  
 str\_detect(parroquia, "^Total|Nacional", negate = T))   
  
pob\_nacional <- pob\_nacional %>%   
 mutate(across(c(provincia, canton, parroquia), str\_to\_upper))

Leemos el acrhvio de codificaciones de las todas las provincias del pasìs, la cual es realizada por el INEC y unimos esta infromación la base de datos censal obtenido en el paso anterior.

dpa\_nacional <- read\_excel("C:/Users/WELCOME/Documents/Curso/Curso Geoestadistica/Examen Final//Data/CODIFICACIÓN\_2022.xlsx", sheet = "CODIGOS")  
pob\_nacional <- pob\_nacional %>% left\_join(dpa\_nacional, by = c("provincia" = "DPA\_DESPRO",  
 "canton" = "DPA\_DESCAN",  
 "parroquia" = "DPA\_DESPAR"))

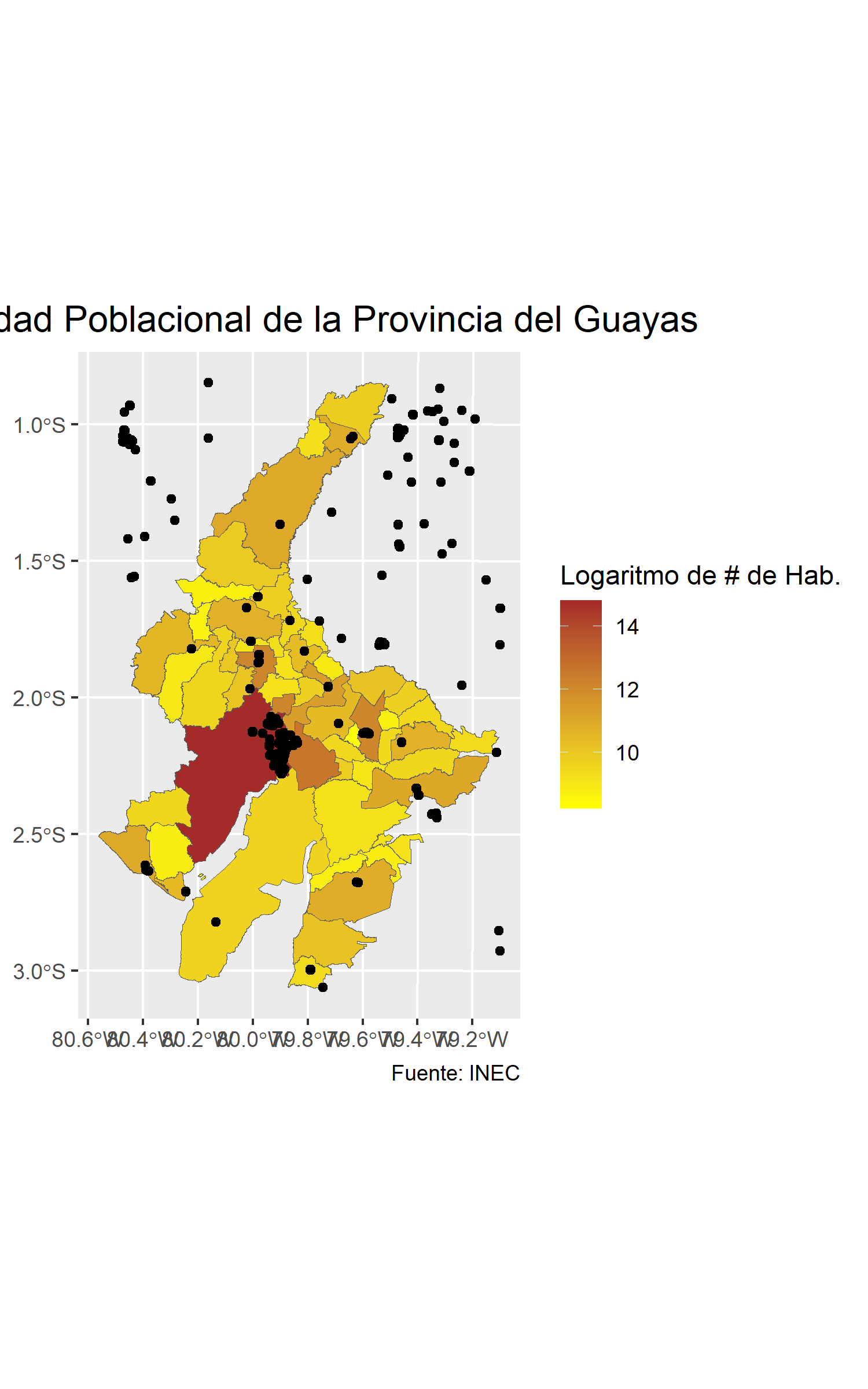
Las zonas obtenidas en la Actividad 1, modificamos la base para poder hacer join con la data censal. Ademas de esto,unimos las parroquias en un shape unico por parroquia

zonas <- zonas %>%   
 mutate(DPA\_PARROQ = str\_sub(zon, 1, 6))  
  
parroquias <- zonas %>% group\_by(DPA\_PARROQ) %>%  
 summarise(Shape = st\_union(Shape))  
  
parroquias <- parroquias %>% left\_join(pob\_nacional)

# Parte 3: Cálculo del Número Total de Amenidades

Vamos este punto, usaremos un mapa de calor , el cual nos ayude a identificar la densidad poblacional de la provincia . Esto lo hacemos con la función ggplot y además , dado que tenemos una distribucìon muy desigual, usamos la función logaritmos para poder tener una visualizacion mucho mas visual.

parroquias %>% ggplot()+  
 geom\_sf(aes(fill = log(total\_nacional)))+  
 geom\_sf(data = puntos)+  
 scale\_fill\_gradient(low = "yellow", high = "brown")+  
 labs(title = "Densidad Poblacional de la Provincia del Guayas",  
 caption = "Fuente: INEC",  
 fill ='Logaritmo de # de Hab.')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



La mayor cantidad de hospitales se encuentra en Guayaquil, la cual la podemos identificar por el color mas oscuro. Ademas , cerca al cantón Guayaquil identificamos 3 parroquias mas las cuales también , sacando la ciudad capital, tambien presentan alta concentración de hospitales. Estas parroquias son Eloy Alfaro, Milagro y Samborondón

# Parte 4: Cálculo del Ratio de Amenidades por Habitante

En este punto vamos a calcular el ratio del número de amenidades seleccionadas en relación con el número de habitantes por cada 100 habitantes en cada parroquia de la provincia del Guayas. Para esto calculamos primero el número de hospitales por cada parroquia.

puntos <- st\_transform(x = puntos, crs = st\_crs(parroquias))  
parroquias\_puntos <- st\_join(x = puntos, y = parroquias, join = st\_within)  
parroquias\_puntos <- parroquias\_puntos %>% count(DPA\_PARROQ, name = "hospitales") %>% as\_tibble()  
parroquias <- parroquias %>% left\_join(parroquias\_puntos)  
parroquias <- parroquias %>% mutate(hospitales = replace\_na(hospitales, 0))

Y vemos un resumen de lop 5 parroquias con mas número de hospitales de la provincia, siendo Guayaquil la que mayor concentración tiene, seguido de Milagro, Eloy Alfaro ,Samborondón y Daule

resumen<-parroquias %>% select(parroquia,hospitales) %>% arrange(desc(hospitales)) %>% as.data.frame()  
head(resumen %>% select(-Shape),5)

parroquia hospitales  
1 GUAYAQUIL 507  
2 MILAGRO 36  
3 ELOY ALFARO 23  
4 SAMBORONDÓN 18  
5 DAULE 14

Una vez , calculado este valor, procedemos al cálculo del indicador mencionado.

parroquias <- parroquias %>% mutate(ratio = (hospitales/total\_nacional)\*1000)  
resumen\_2<-parroquias %>% select(parroquia,ratio) %>% arrange(desc(ratio)) %>% as.data.frame() %>% head(5)  
  
resumen\_3<-parroquias %>% filter(ratio!=0) %>%select(parroquia,ratio) %>% arrange(ratio) %>% as.data.frame() %>% head(5)  
  
resumen\_4<-parroquias %>% filter(ratio==0)

A continuacion vemos las 5 mejores parroquias con respecto al indicador.

resumen\_2 %>%select(-Shape)

parroquia ratio  
1 LIMONAL 0.6373486  
2 TENGUEL 0.4152134  
3 VELASCO IBARRA 0.2355204  
4 PALESTINA 0.2219879  
5 SAMBORONDÓN 0.2199037

Además podemos indicar que existen 32 parroquias que no tienen ningun hospital disponible,

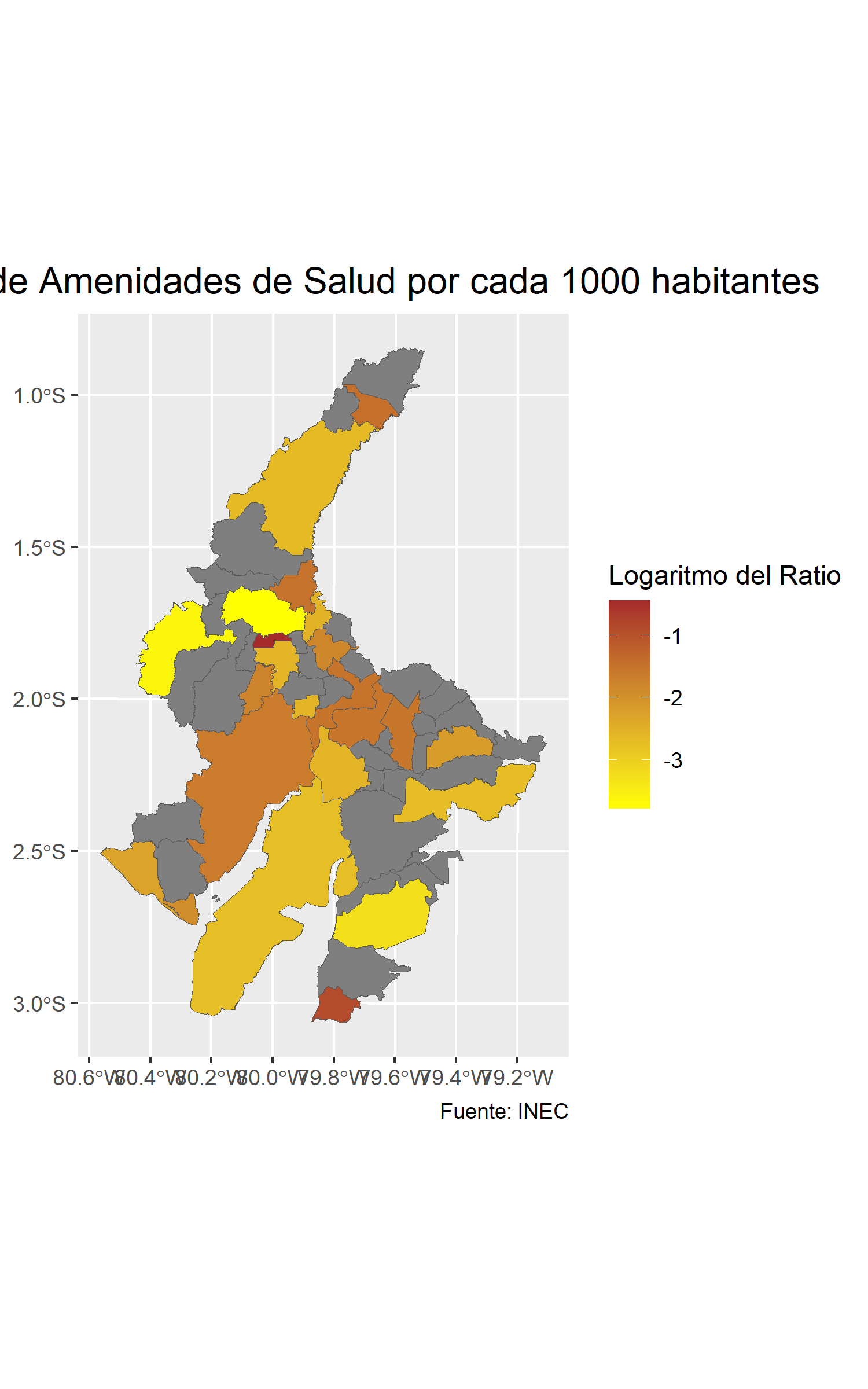
Y ahora vemos el top 5 de las peores parroquias con respecto a este indicador, sacando todas las parroquias que no tienen ningún hospital.

resumen\_3 %>%select(-Shape)

parroquia ratio  
1 SANTA LUCÍA 0.02288330  
2 PEDRO CARBO 0.02603828  
3 NARANJAL 0.03846006  
4 PUNÁ 0.06546645  
5 EL TRIUNFO 0.06607093

Las parroquias que apreciamos sin color en donde no existe presencia de hospitales. Esto explica porque hay algunas parroquias tienen alto valor de indicador ya que la población de estas parroquias deben trasladarse a otras parroquias para poder ser atendidos.

parroquias %>% ggplot()+  
 geom\_sf(aes(fill = log(ratio)))+  
 scale\_fill\_gradient(low = "yellow", high = "brown")+  
 labs(title = "Indicador de Amenidades de Salud por cada 1000 habitantes",  
 caption = "Fuente: INEC",fill ='Logaritmo del Ratio',  
 fill ='Logaritmo de Numero de Pob.')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



# Parte 5: Cálculo del Índice de Moran para el Indicador de Amenidades de

Salud

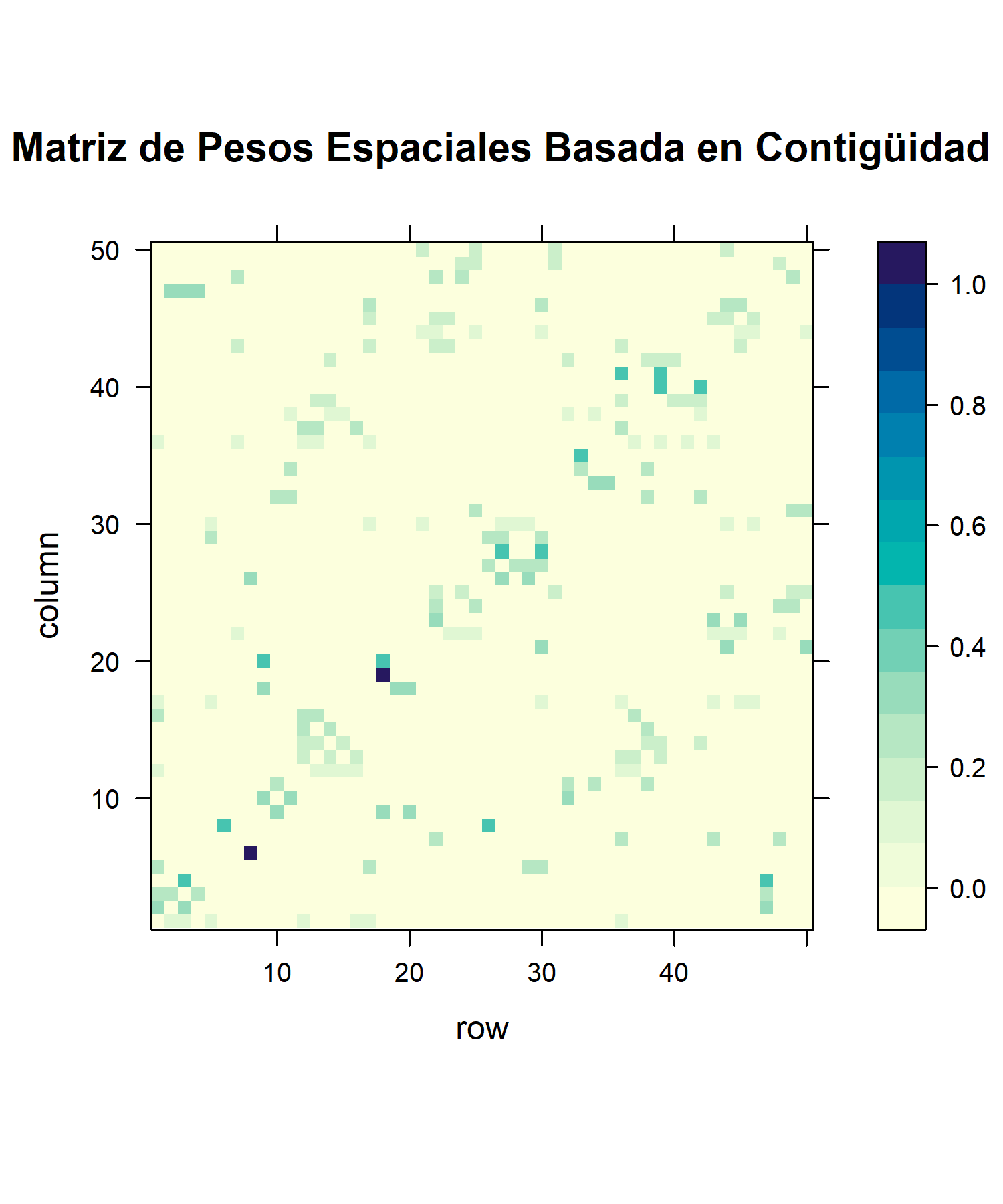
En este punto vamos a calcular el Índice de Moran para el indicador de número de amenidades de salud por cada 100 habitantes en las parroquias de una provincia. Este índice nos permitirá evaluar si existe autocorrelación espacial en la distribución de las amenidades de salud, lo cual es clave para entender la distribución y concentración de estos servicios en diferentes áreas.

Calculo de la matriz de Vecindad binaria:

vecinos <- poly2nb(parroquias, queen = T)  
vecinos\_pesos <- nb2listw(vecinos, style = "W")  
matriz <- listw2mat(vecinos\_pesos)

Cálculo de Matriz de Pesos Espaciales de Contigüidad:

levelplot(t(matriz[1:50,1:50]),  
 scales = list(y = list(at = seq(10, 50, by = 10),  
 labels = seq(10, 50, by = 10))),  
 main = "Matriz de Pesos Espaciales Basada en Contigüidad")



Cálculo del Indice de Morán:

moran <- moran.test(parroquias$hospitales, vecinos\_pesos, alternative = "two.sided")  
  
moran

Moran I test under randomisation  
  
data: parroquias$hospitales   
weights: vecinos\_pesos   
  
Moran I statistic standard deviate = -0.42806, p-value = 0.6686  
alternative hypothesis: two.sided  
sample estimates:  
Moran I statistic Expectation Variance   
 -0.0255997093 -0.0188679245 0.0002473183

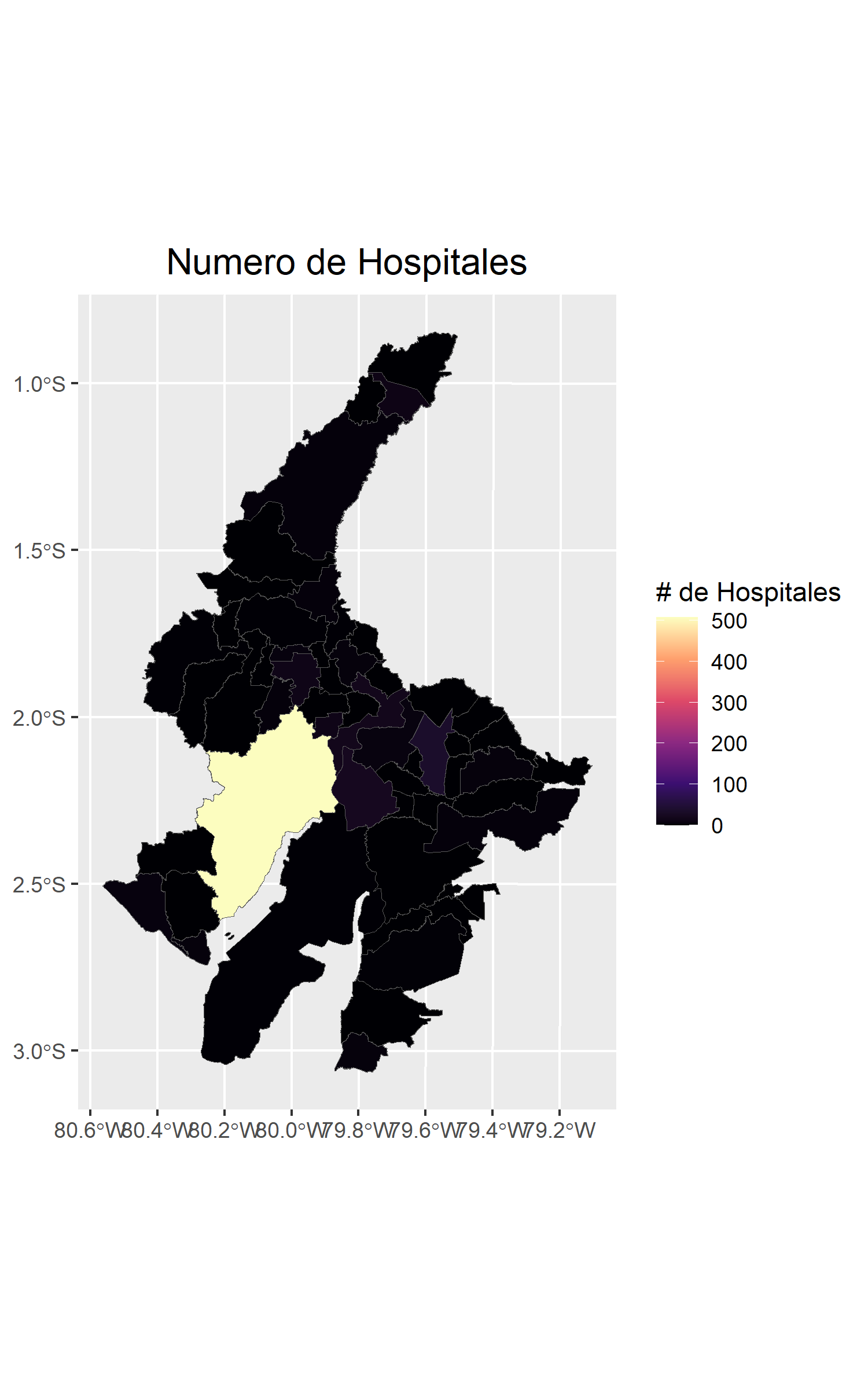
De acuerdo al Índice de Moran, no existe autocorrelación espacial entre el número de hospitales dentro de la provincia de Guayas al tener un p-valor de 0.6686; el cual es superior al de 0.05.

Cálculo del indice local de Moran y composición de clusteres:

local\_moran<- localmoran(parroquias$hospitales, listw = vecinos\_pesos)  
  
#head(local\_moran)  
  
clusters <- attributes(local\_moran)$quadr  
parroquias <- bind\_cols(parroquias, clusters)  
local\_moran <- local\_moran %>% as.tibble()  
  
parroquias <- parroquias %>% bind\_cols(local\_moran)

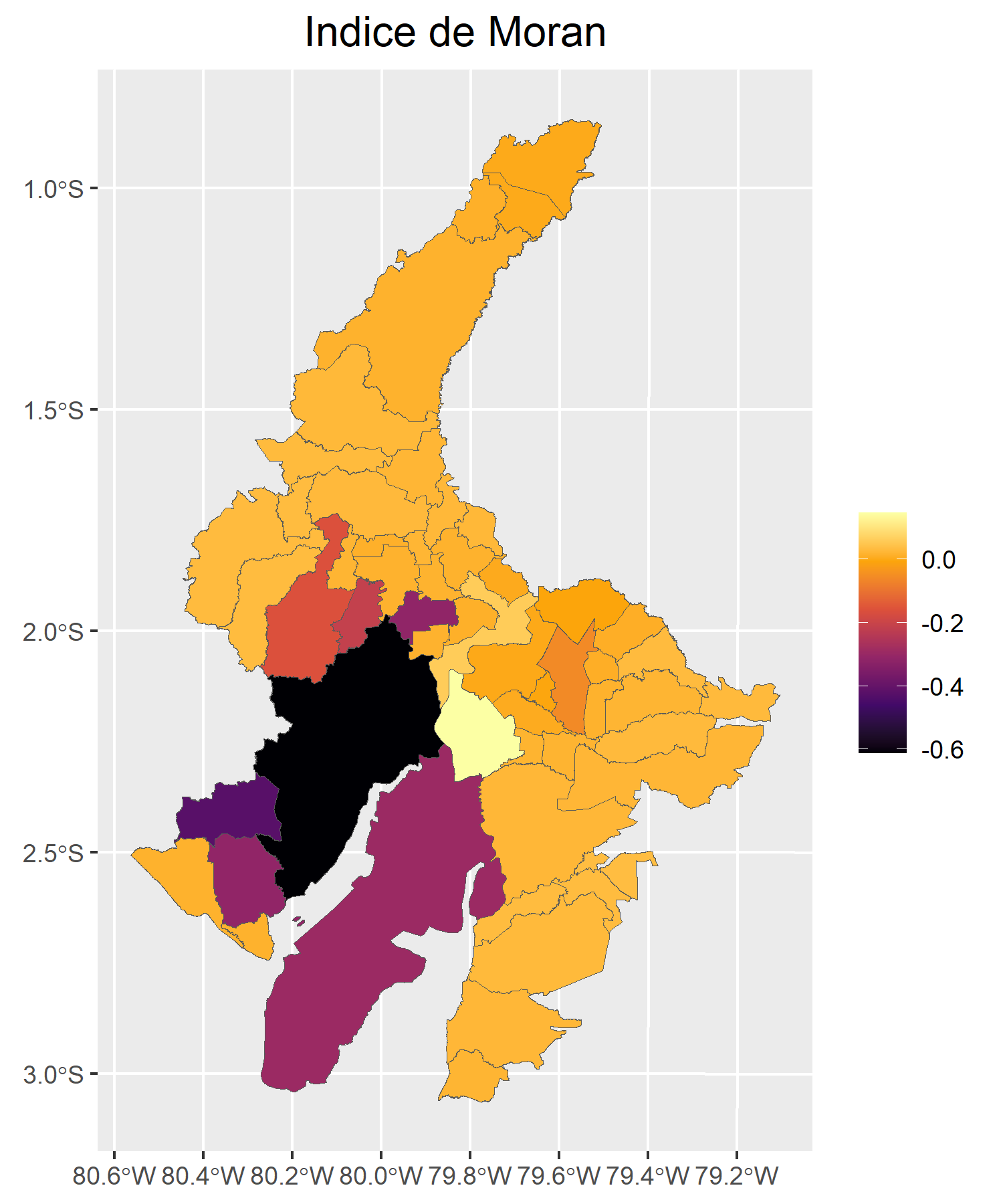
### 1. Mapa del Número Promedio de Hospitales

parroquias %>%   
 ggplot() +   
 geom\_sf(aes(fill = hospitales)) +  
 scale\_fill\_viridis\_c(option = "A") + labs(title='Numero de Hospitales',fill='# de Hospitales')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



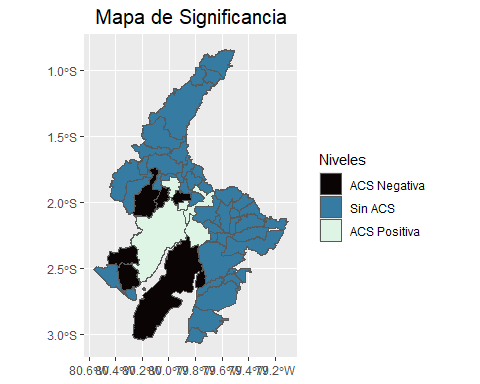
### 2. Mapa del Índice de Moran Local

parroquias %>%   
 ggplot() +   
 geom\_sf(aes(fill = as.numeric(Ii))) +  
 scale\_fill\_viridis\_c(option = "B")+  
 labs(title = "Indice de Moran",fill='')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



### 3. Mapa de Significancia del Índice de Moran Local

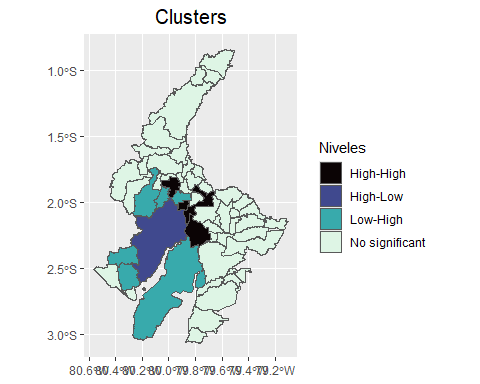
parroquias %>%   
 mutate(Z.Ii = as.numeric(Z.Ii),  
 Z.Ii = cut(Z.Ii, breaks = c(-Inf, -1.65, 1.65, Inf), labels = c("ACS Negativa", "Sin ACS", "ACS Positiva"))) %>%   
 ggplot() +   
 geom\_sf(aes(fill = Z.Ii)) +  
 scale\_fill\_viridis\_d(option = "G")+labs(title='Mapa de Significancia',fill='Niveles')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



Dentro de la provincia, se puede ver que casi en su totalidad no poseen autocorrelación; mientras que la parroquia Guayaquil, Eloy Alfaro, Samborondon y Daule tiene autocorrelación positiva y sus parroquias aledañas tienen autocorrelación negativa.

### 4. Clústers

parroquias %>%   
 mutate(mean = if\_else(between(as.numeric(Z.Ii), -1.65, 1.65),"No significant",mean)) %>%   
 ggplot() +   
 geom\_sf(aes(fill = mean)) +  
 scale\_fill\_viridis\_d(option = "G")+labs(title='Clusters',fill='Niveles')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



En la parroquia de Guayaquil existen clusteres de relación Alta- Baja. Existen 5 cantones que son relacion Alta- Alta. 6 cantones relación Baja- Alta y los demás parroquias sin significancia estadìstica.

# Parte 6: Actividad opcional, Análisis de Buffers y Reflexión sobre los Límites de OpenStreetMap

### 1. Crear un Buffer alrededor del Polígono de Diners Club

Ubicamos nuestra zona en la provincia de Pichincha, parroquia Iñaquito

zonas <- st\_read("C:/Users/WELCOME/Documents/Curso/Curso Geoestadistica/Examen Final/Data/GEODATABASE\_NACIONAL\_2021/GEODATABASE\_NACIONAL\_2021.gdb/",   
 layer = "zon\_a")

Reading layer `zon\_a' from data source   
 `C:\Users\WELCOME\Documents\Curso\Curso Geoestadistica\Examen Final\Data\GEODATABASE\_NACIONAL\_2021\GEODATABASE\_NACIONAL\_2021.gdb'   
 using driver `OpenFileGDB'  
Simple feature collection with 5888 features and 6 fields  
Geometry type: MULTIPOLYGON  
Dimension: XY  
Bounding box: xmin: -685704.9 ymin: 9445216 xmax: 1147852 ymax: 10162550  
Projected CRS: WGS 84 / UTM zone 17S

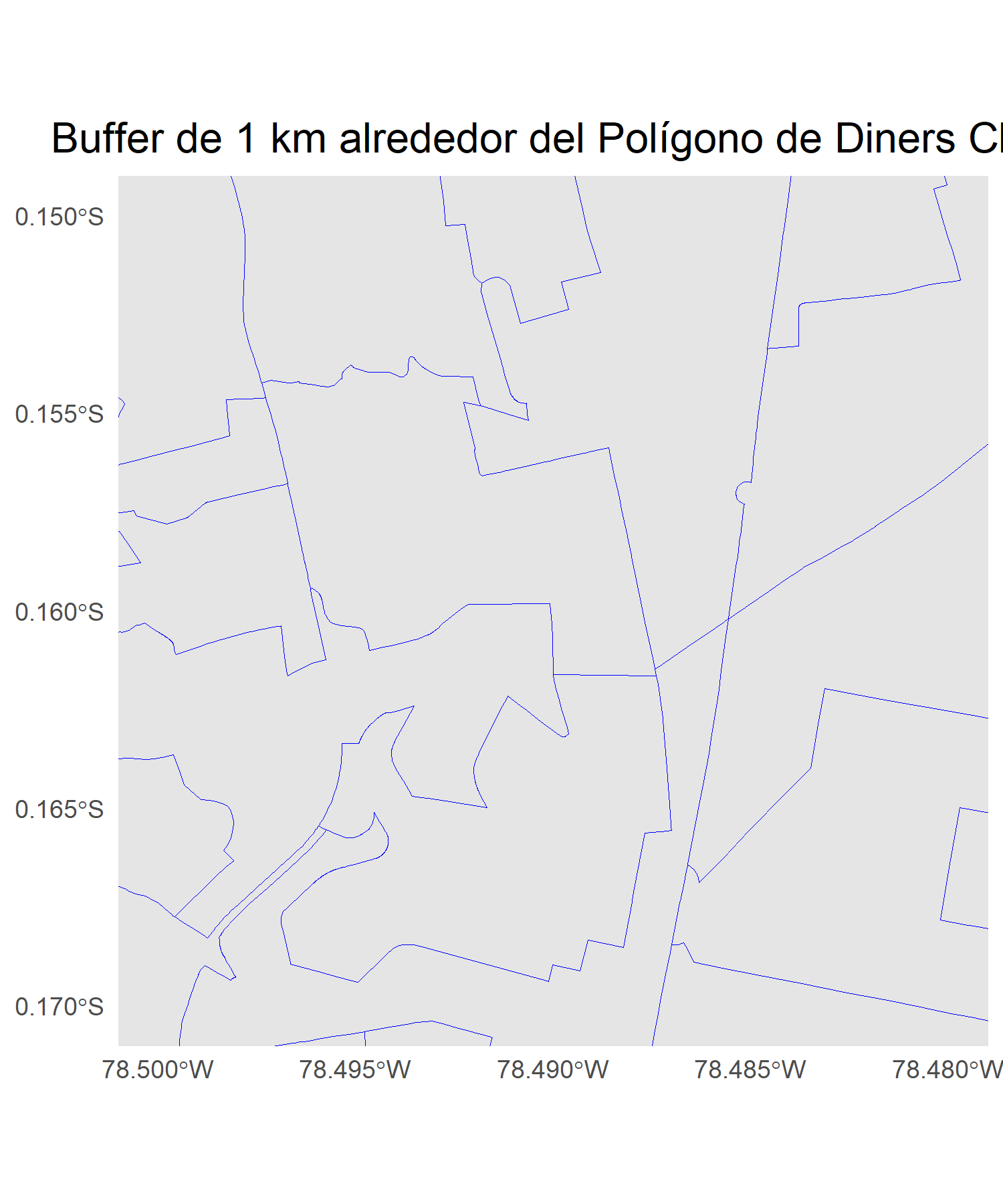
dmq\_zonas <- zonas %>% filter(str\_detect(zon, "^170150"))  
  
# Crear un objeto de ejemplo   
zoom\_to <- "170150192"  
  
poli <- dmq\_zonas %>% filter(zon == zoom\_to)  
  
# Crear un buffer de 500 m alrededor del polígono  
buffer <- st\_buffer(poli, dist = 500)  
  
class(buffer$Shape)

[1] "sfc\_POLYGON" "sfc"

# Definir los límites del mapa para el zoom  
lon\_bounds <- c(-78.480, -78.500) # Ajustar según sea necesario  
lat\_bounds <- c(-0.17, -0.15)  
  
  
# Transformar las coordenadas de las zonas para la visualización  
dmq\_4326 <- st\_transform(dmq\_zonas, crs = 4326)  
  
# Mostrar el buffer en un mapa  
  
class(poli$Shape)

[1] "sfc\_MULTIPOLYGON" "sfc"

ggplot() +  
 geom\_sf(data = dmq\_4326, color = "blue", size = 1) +  
 geom\_sf(data = buffer, fill = "grey", alpha = 0.5) +  
 theme\_minimal() +  
 ggtitle("Buffer de 1 km alrededor del Polígono de Diners Club") +  
 coord\_sf(xlim = lon\_bounds, ylim = lat\_bounds)+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



### 2. Aplicar el Buffer al Centroide del Polígono

Aplicaremos el buffer al centroide del polígono en lugar de al polígono completo.

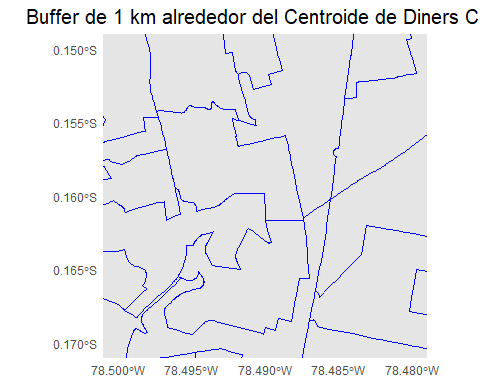
# Calcular el centroide del polígono de Diners Club  
centroid <- st\_centroid(poli)

Warning: st\_centroid assumes attributes are constant over geometries

# Crear un buffer de 1 km alrededor del centroide  
buffer\_centroid <- st\_buffer(centroid, dist = 500)  
  
class(centroid$Shape)

[1] "sfc\_POINT" "sfc"

# Mostrar el buffer del centroide en un mapa  
ggplot() +  
 geom\_sf(data = dmq\_4326, color = "blue", size = 1) +  
 geom\_sf(data = buffer\_centroid, fill = "grey", alpha = 0.5) +  
 theme\_minimal() +  
 ggtitle("Buffer de 1 km alrededor del Centroide de Diners Club") +  
 coord\_sf(xlim = lon\_bounds, ylim = lat\_bounds)+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



### 3. Buscar Hospitales dentro de los Buffers

Ahora vamos a identificar las zonas que contienen tanto restaurantes como hospitales dentro de los buffers de 500 metros.

library(osmdata)  
  
bbox <- getbb("Quito, Ecuador")  
  
hospitales <- opq(bbox = bbox) %>%   
 add\_osm\_feature(key = "amenity",value = "hospital") %>%   
 osmdata\_sf()  
  
hospital\_points <- hospitales$osm\_points  
  
# Transformar las coordenadas de los puntos al mismo CRS que `dmq\_zonas`  
hospitals\_points <- st\_transform(hospital\_points, crs = st\_crs(dmq\_zonas))  
  
# Crear buffers de 1 km alrededor de cada zona censal  
dmq\_zonas\_buffers <- st\_buffer(st\_centroid(dmq\_zonas), dist = 500)  
  
# Identificar zonas con hospitales en el buffer  
zones\_with\_hospitals <- st\_join(hospitals\_points, dmq\_zonas\_buffers,   
 join = st\_within)

### 4. Crear Listas de IDs que Tienen Ambas Amenidades

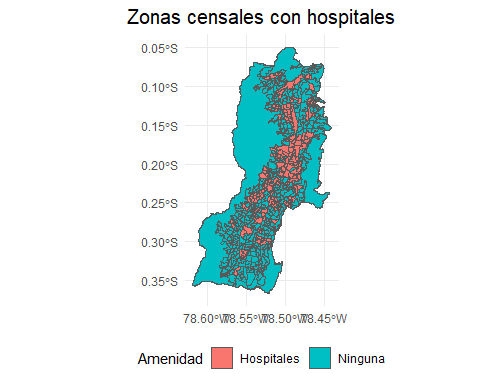
Creamos listas que identifican las zonas que tienen un hospital dentro del buffer.

amenidades <- list(hospitales = zones\_with\_hospitals) %>%  
 map(~.x %>% as\_tibble() %>% distinct(zon))  
  
ambos <- amenidades %>%  
 reduce(inner\_join, by = "zon")  
  
amenidades <- append(amenidades, list(ambos = ambos))

### 5. Mostrar el Mapa Completo con Amenidades

Mostramos un mapa completo que indica las zonas que tienen hospitales O ninguna amenidad.

dmq\_zonas <- dmq\_zonas %>%  
 mutate(  
 amenidad = case\_when(  
 zon %in% amenidades$hospitales$zon ~ "Hospitales",  
 TRUE ~ "Ninguna"  
 )  
 )  
  
dmq\_zonas %>%   
ggplot() +  
 geom\_sf(aes(fill = amenidad)) +  
 theme\_minimal() +  
 theme(legend.position = "bottom") +  
 labs(title = "Zonas censales con hospitales", fill = "Amenidad")+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))



### 6. Zoom sobre Diners Club y sus Amenidades

Hacemos zoom sobre el polígono de Diners Club para mostrar cómo se ven las amenidades en esta zona específica.

# Definir los límites del mapa para el zoom  
lon\_bounds <- c(-78.480, -78.500) # Ajustar según sea necesario  
lat\_bounds <- c(-0.17, -0.15)  
  
# Transformar las coordenadas de las zonas para la visualización  
dmq <- st\_transform(dmq\_zonas, crs = 4326)  
  
buffer\_poli\_2 <- st\_buffer(st\_centroid(poli),dist = 500)

Warning: st\_centroid assumes attributes are constant over geometries

# Mostrar el mapa con zoom sobre la Politécnica  
ggplot(dmq) +  
 geom\_sf(aes(fill = amenidad)) +  
 geom\_sf(data = buffer\_poli\_2, fill = "grey", alpha = 0.5) +  
 coord\_sf(xlim = lon\_bounds, ylim = lat\_bounds) +  
 theme\_minimal() +  
 ggtitle("Zoom sobre Diners Club y sus Amenidades")+labs(fill='Amenidad')+  
 theme(plot.title = element\_text(hjust = 0.5,size=15))

