

Análisis de la distribucion territorial en República Dominicana de personas con dificultad para caminar o subir escalones y su relación de dependencia con niveles educativos alcanzados.

Análisis de la region con los mayores niveles de precipitación ocurrida en el año 1982 a partir de datos suministrados por la oficina nacional de meteorología para la República Dominicana.

**Wanda Lisselote Binet y Ramon Correa**    *Estudiantes de la Maestría en Teledetección y Ciencias de la información Geográfica, Universidad Autónoma de Santo Domingo (UASD) - Módulo de Análisis Espacial*

---

Mi resumen

*Keywords:* Dificultad Caminar, Dificultad subir escalones, Precipitación, Nivel educativo alcanzado

---

## 1 Introducción

Estos procesos de análisis utilizando R, se realizan como trabajo final del módulo de Análisis Espacial de la Maestría en teledetección y Ciencias de la información Geográfica. Su objetivo específico es realizar procesos de análisis espacial, tomando en cuenta los datos del Censo Nacional de población y vivienda del 2010, que nos ayuden a definir la distribución espacial de las personas que tienen dificultad para caminar o subir escalones en el país, localizando los puntos de concentración de dicha población por municipio. De igual forma se hará la modelización de la relación entre esta población con los niveles educativos alcanzados.

De igual manera, mediante el uso de la interpolación de kriging, se efectúa el análisis de los datos de precipitación obtenidos de la oficina nacional de meteorología del año 1982 para conocer los municipios del país donde se evidencia una mayor cantidad de lluvia durante ese periodo.

## 2 Información de soporte

Se utiliza como información de soporte las bases de datos suministradas por el profesor Martínez Batlle que está localizada en el directorio data: \* Capa de división política de municipios del Censo Nacional de Población y Vivienda 2010, localizada en el Geopackage divisionRD.gpkg. \* Datos de precipitación anual, localizados en los archivos onamet\_prec\_anual\_sf. \* Base de datos estadísticos de la Oficina Nacional de Estadística, localizada en el archivo vivpersgeom\_sf.RDS.

### 3 Referencias

- Se utiliza como material de apoyo los scripts practicados durante las sesiones de trabajo en aula con el profesor José Ramón Martínez Batlle, durante el Módulo Análisis Espacial (Introducción a R, Simple features y análisis exploratorio de datos espaciales (ESDA), Vecindad, Autocorrelación, Datos puntuales-Geoestadística, Modelización de datos espaciales basados en geometrías poligonales).
- <https://desktop.arcgis.com/es/arcmap/latest/tools/spatial-statistics-toolbox/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm>
- <https://medium.com/high-data/el-concepto-de-heterocedasticidad-36cda43bb8f7>
- <http://academic.uprm.edu/eacuna/miniman9sl.pdf>
- [https://es.wikipedia.org/wiki/Modelo\\_autorregresivo](https://es.wikipedia.org/wiki/Modelo_autorregresivo)
- <http://www.cartagena99.com/recursos/alumnos/apuntes/Tema%202%20-%20Regresion%20lineal.pdf>

### 4 Script reproducible

#### 5 Carga de librerías de R a memoria

```
library(sf)
## Linking to GEOS 3.7.1, GDAL 2.4.2, PROJ 5.2.0
library(sp)
library(tidyverse)
## -- Attaching packages ----- tidyverse
## v ggplot2 3.2.1      v purrr   0.3.3
## v tibble  2.1.3      v dplyr    0.8.3
## v tidyverse 1.0.0     v stringr  1.4.0
## v readr   1.3.1      vforcats  0.4.0
## -- Conflicts ----- tidyverse_conflict
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
library(spdep)
## Loading required package: spData
## To access larger datasets in this package, install the spDataLarge
## package with: 'install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source)'
library(lmtest)
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
library(tmap)
library(RColorBrewer)
library(gstat)
## Registered S3 method overwritten by 'xts':
##   method      from
```

```

##   as.zoo.xts zoo
library(stars)
## Loading required package: abind
source('data/lisaclusters.R')

```

## 6 FASE 1 - ANALISIS DE AUTOCORRELACION ESPACIAL

### 6.1 Metodología

Para arribar a los resultados buscados primeramente se han cargado en memoria el archivos de datos estadísticos del IX censo nacional del 2010 y el geopackage contentivo de la base espaciales con la división política municipal. Procediendo posteriormente a realizar un análisis exploratorio de los datos para informaciones estadísticas básicas, histogramas, y pruebas para comprobar la existencia de una distribución normal de los datos.

Confirmado esto, se procede a crear objetos de vecindad por contiguidad y los ponderadores espaciales u objeto de pesos, tanto el estilo weighted como el binario. Para entonces iniciar los procesos para comprobar la autocorrelación espacial, haciendo la prueba de Breuch-Pagan, de homocedasticidad de la variable transformada (es decir, que la varianza de los errores de la variable es constante a lo largo del tiempo), el Test I de Moran global (que mide la autocorrelación espacial basada en las ubicaciones) y el Test de I de Moran local (que identifica clusters con valores altos, bajos y también valores atípicos espaciales), para concluir con la creación del grafo Lisa Cluster que nos muestra el comportamiento de la variable estudiada: personas que tienen dificultad para caminar o subir escalones en el país.

### 6.2 Carga de los archivos espaciales y de datos a utilizar.

```

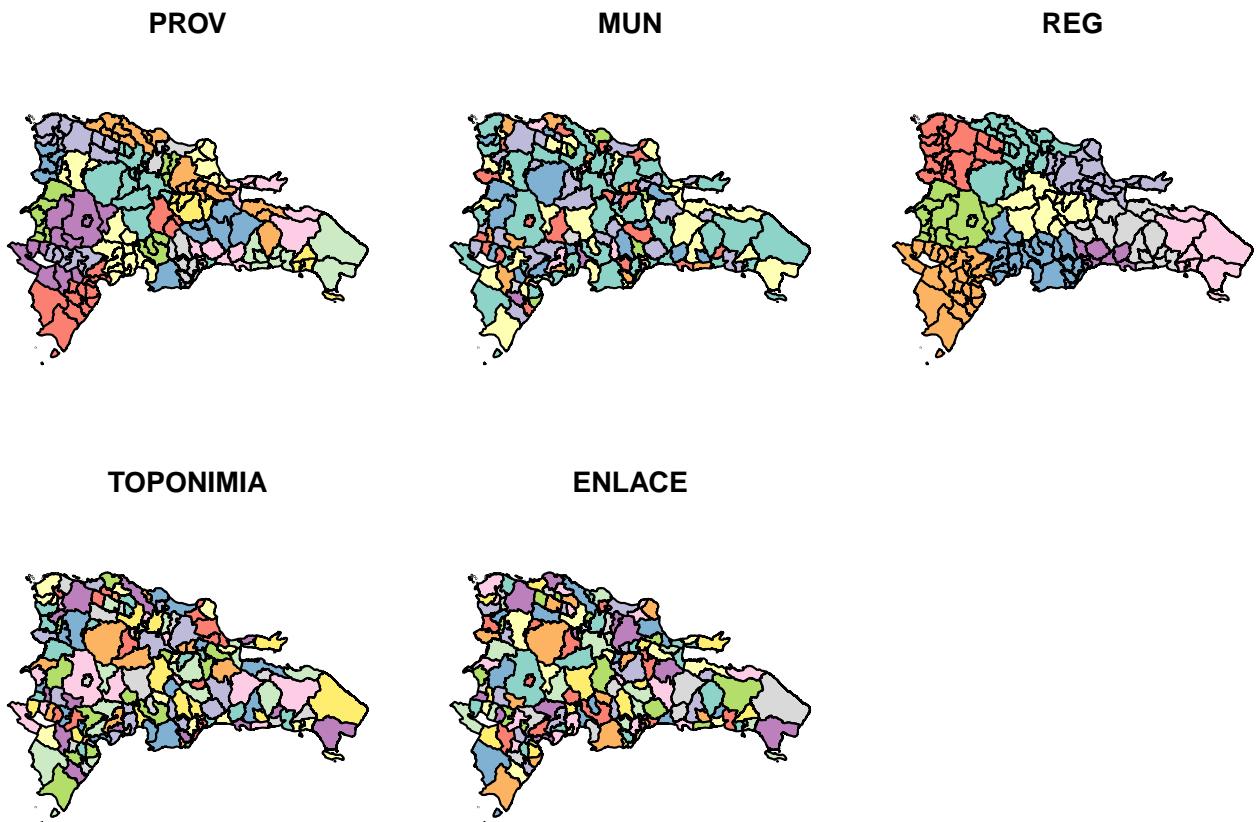
# Carga en memoria del geopackage de division territorial
st_layers('data/divisionRD.gpkg')
## Driver: GPKG
## Available layers:
##   layer_name geometry_type features fields
## 1 PROVCenso2010      Polygon      32      4
## 2 REGCenso2010      Polygon      10      2
## 3 MUNCenso2010      Polygon    155      5
muni.sf <- st_read(dsn = 'data/divisionRD.gpkg', layer = 'MUNCenso2010', quiet=T)
muni.sf
## Simple feature collection with 155 features and 5 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:            xmin: 182215.8 ymin: 1933512 xmax: 571429.3 ymax: 2205216
## epsg (SRID):    32619
## proj4string:    +proj=utm +zone=19 +datum=WGS84 +units=m +no_defs
## First 10 features:
##   PROV MUN REG          TOPONIMIA ENLACE
## 1    01  01  10 SANTO DOMINGO DE GUZMÁN 100101
## 2    02  01  05           AZUA 050201

```

```

## 3 02 02 05 LAS CHARCAS 050202
## 4 02 03 05 LAS YAYAS DE VIAJAMA 050203
## 5 02 04 05 PADRE LAS CASAS 050204
## 6 02 05 05 PERALTA 050205
## 7 02 06 05 SABANA YEGUA 050206
## 8 02 07 05 PUEBLO VIEJO 050207
## 9 02 08 05 TÁBARA ARRIBA 050208
## 10 02 09 05 GUAYABAL 050209
##
## geom
## 1 MULTIPOLYGON (((397122.7 20...
## 2 MULTIPOLYGON (((318830.5 20...
## 3 MULTIPOLYGON (((333950.3 20...
## 4 MULTIPOLYGON (((288107.3 20...
## 5 MULTIPOLYGON (((303931.9 21...
## 6 MULTIPOLYGON (((307770.9 20...
## 7 MULTIPOLYGON (((304091.3 20...
## 8 MULTIPOLYGON (((314474 2040...
## 9 MULTIPOLYGON (((302204.4 20...
## 10 MULTIPOLYGON (((313520.7 20...
summary(muni.sf)
##          PROV        MUN       REG      TOPONIMIA
## 04       : 11     01     :32    06     :24 ALTAMIRA   : 1
## 02       : 10     02     :31    05     :23 ARENOSO   : 1
## 18       :  9     03     :27    01     :22 AZUA       : 1
## 25       :  9     04     :19    03     :17 BAJOS DE HAINA: 1
## 21       :  8     05     :15    04     :17 BANÍ       : 1
## 06       :  7     06     :12    09     :14 BÁNICA    : 1
## (Other):101 (Other):19 (Other):38 (Other)   :149
##
## ENLACE           geom
## 010901 : 1 MULTIPOLYGON :155
## 010902 : 1 epsg:32619   : 0
## 010903 : 1 +proj=utm ...: 0
## 010904 : 1
## 011801 : 1
## 011802 : 1
## (Other):149
plot(muni.sf)

```



```
# Carga los datos del censo 2010
vivpersgeom_sf <- readRDS('data/vivpersgeom_sf.RDS')
censo <- vivpersgeom_sf %>% select(matches('ENLACE|TOPONIMIA|Población total|Condición de ocupación'))

censo <- censo %>% mutate("Cantviv"='Condición de ocupación: Ocupada con personas presentes'+ 'Ocupada sin personas presentes')

censo <- censo %>% mutate("PorcPersD"= 'Dificultad para Caminar o subir escalones: Si'/'Población total')

censo <- censo %>% mutate("PorcPersD_log"= log(censo$PorcPersD))

structure(censo)
## Simple feature collection with 155 features and 14 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:            xmin: -72.01147 ymin: 17.47033 xmax: -68.32354 ymax: 19.93211
## epsg (SRID):    4326
## proj4string:    +proj=longlat +datum=WGS84 +no_defs
## First 10 features:
##           TOPONIMIA ENLACE
## 1  SANTO DOMINGO DE GUZMÁN 100101
## 2                  AZUA 050201
## 3          LAS CHARCAS 050202
## 4  LAS YAYAS DE VIAJAMA 050203
```

## 5	PADRE LAS CASAS	050204
## 6	PERALTA	050205
## 7	SABANA YEGUA	050206
## 8	PUEBLO VIEJO	050207
## 9	TÁBARA ARRIBA	050208
## 10	GUAYABAL	050209
##	Condición de ocupación: Ocupada con personas presentes	
## 1		288362
## 2		22797
## 3		3147
## 4		4807
## 5		5243
## 6		3144
## 7		4828
## 8		2689
## 9		4108
## 10		1386
##	Condición de ocupación: Desocupada Población total	
## 1	42200	965040
## 2	1920	91345
## 3	944	11243
## 4	798	17620
## 5	1345	20041
## 6	446	15257
## 7	594	19020
## 8	208	11235
## 9	426	17647
## 10	464	5263
##	Dificultad para Caminar o subir escalones: Si	
## 1		31271
## 2		2616
## 3		323
## 4		770
## 5		804
## 6		320
## 7		675
## 8		266
## 9		437
## 10		234
##	Asiste o asistió a la escuela: Nunca asistió	
## 1		45138
## 2		12617
## 3		2329
## 4		3442
## 5		4336
## 6		3528
## 7		2813

```

## 8 1629
## 9 3860
## 10 909
## Nivel educativo más alto al que asistió: Preprimaria
## 1 73199
## 2 7718
## 3 965
## 4 1400
## 5 2027
## 6 1360
## 7 1079
## 8 961
## 9 1344
## 10 240
## Nivel educativo más alto al que asistió: Primaria o básica
## 1 290061
## 2 37701
## 3 5361
## 4 8439
## 5 8458
## 6 5736
## 7 9246
## 8 5270
## 9 7916
## 10 2458
## Nivel educativo más alto al que asistió: Secundaria o media
## 1 248102
## 2 19003
## 3 1616
## 4 2757
## 5 3104
## 6 2823
## 7 3676
## 8 2133
## 9 2869
## 10 924
## Nivel educativo más alto al que asistió: Universitaria o superior
## 1 257590
## 2 8601
## 3 204
## 4 541
## 5 1075
## 6 1020
## 7 1048
## 8 484
## 9 602
## 10 479

```

```

##                                     geom  Cantviv PorcPersD PorcPersD_log
## 1  MULTIPOLYGON (((-69.89794 1... 330562  3.240384    1.1756918
## 2  MULTIPOLYGON (((-70.71457 1... 24717   2.863868    1.0521731
## 3  MULTIPOLYGON (((-70.50185 1... 4091   2.872899    1.0553215
## 4  MULTIPOLYGON (((-70.85774 1... 5605   4.370034    1.4747708
## 5  MULTIPOLYGON (((-70.77551 1... 6588   4.011776    1.3892340
## 6  MULTIPOLYGON (((-70.73131 1... 3590   2.097398    0.7406975
## 7  MULTIPOLYGON (((-70.83014 1... 5422   3.548896    1.2666365
## 8  MULTIPOLYGON (((-70.79387 1... 2897   2.367601    0.8618773
## 9  MULTIPOLYGON (((-70.83352 1... 4534   2.476342    0.9067823
## 10 MULTIPOLYGON (((-70.68664 1... 1850   4.446133    1.4920348

# Aseguro se trabaja con el mismo sistema de coordenadas
censoT = st_transform(censo, crs=32619)

# Aseguro la relacion entre los archivos con el campo ENLACE y TOPONIMIA
match(censoT$ENLACE, muni.sf$ENLACE)
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
## [18] 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
## [35] 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
## [52] 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68
## [69] 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85
## [86] 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102
## [103] 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119
## [120] 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136
## [137] 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153
## [154] 154 155

match(censoT$TOPONIMIA, muni.sf$TOPONIMIA)
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
## [18] 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
## [35] 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
## [52] 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68
## [69] 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85
## [86] 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102
## [103] 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119
## [120] 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136
## [137] 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153
## [154] 154 155

```

### 6.3 Análisis Exploratorio ESDA

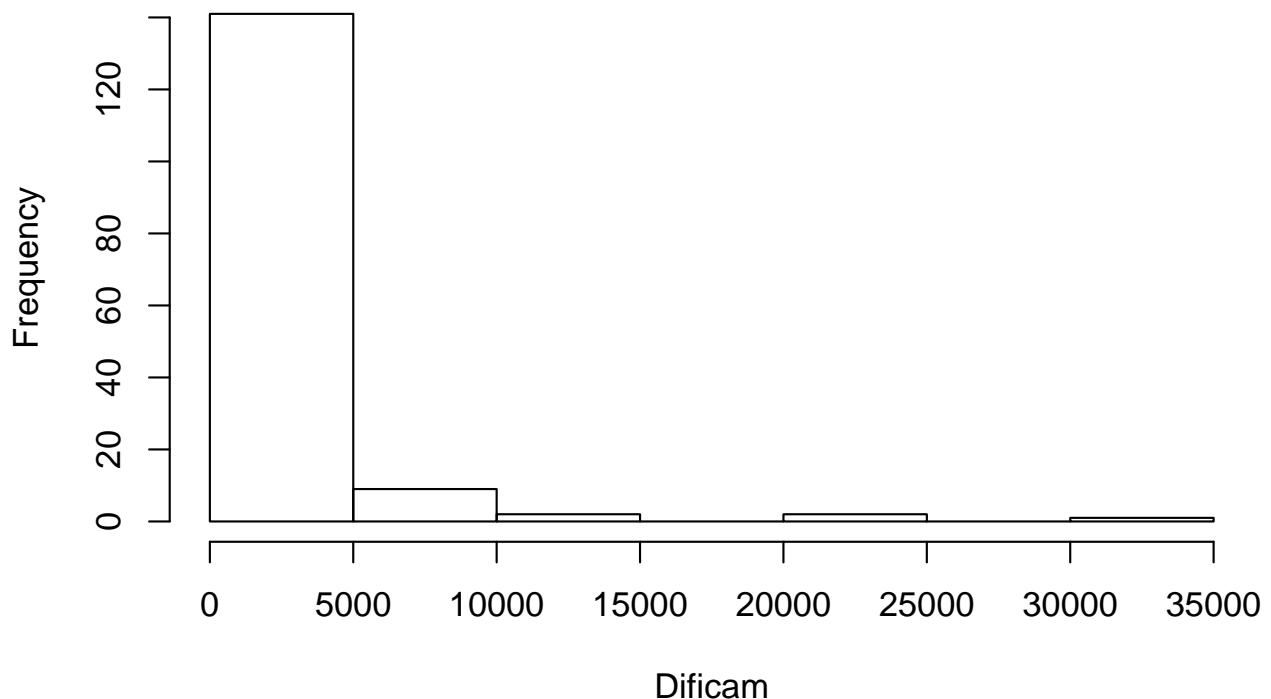
```

nrow(censoT)
## [1] 155
summary(censoT$"Dificultad para Caminar o subir escalones: Si")
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.

```

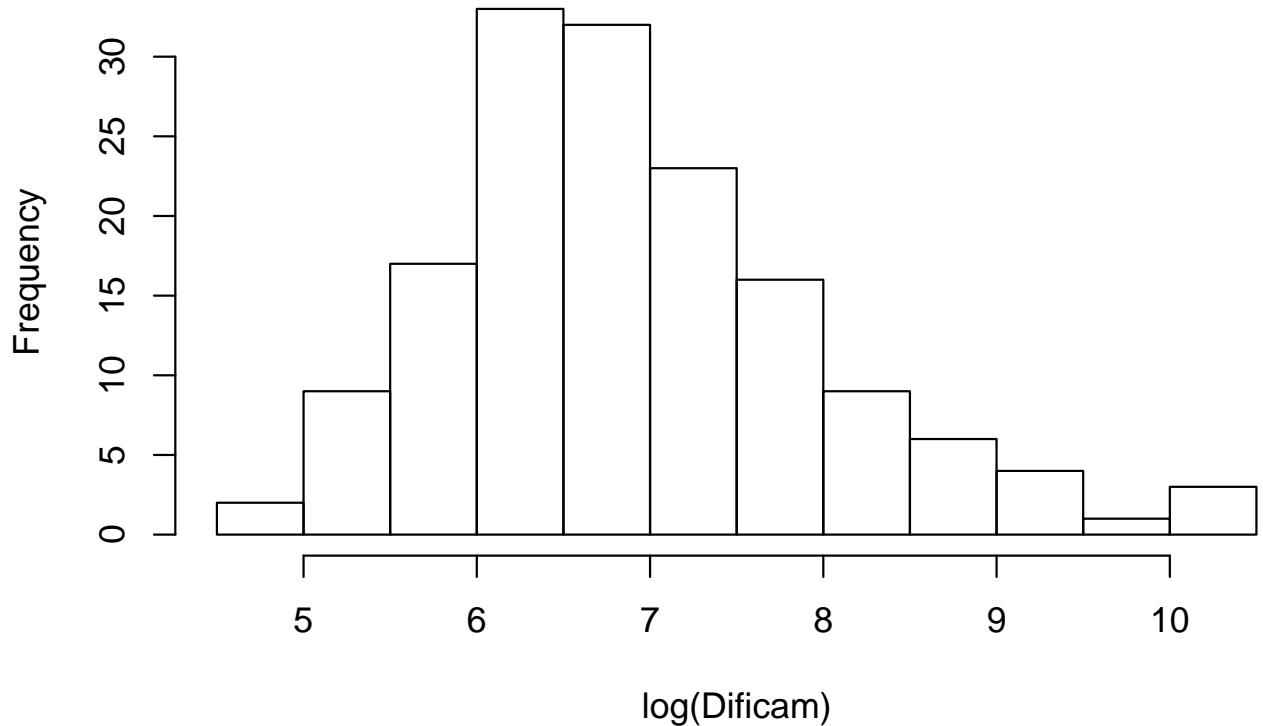
```
##      113     443     902    2114    1819   31271
Dificam <- censoT$`Dificultad para Caminar o subir escalones: Si`
hist(Dificam)
```

**Histogram of Dificam**



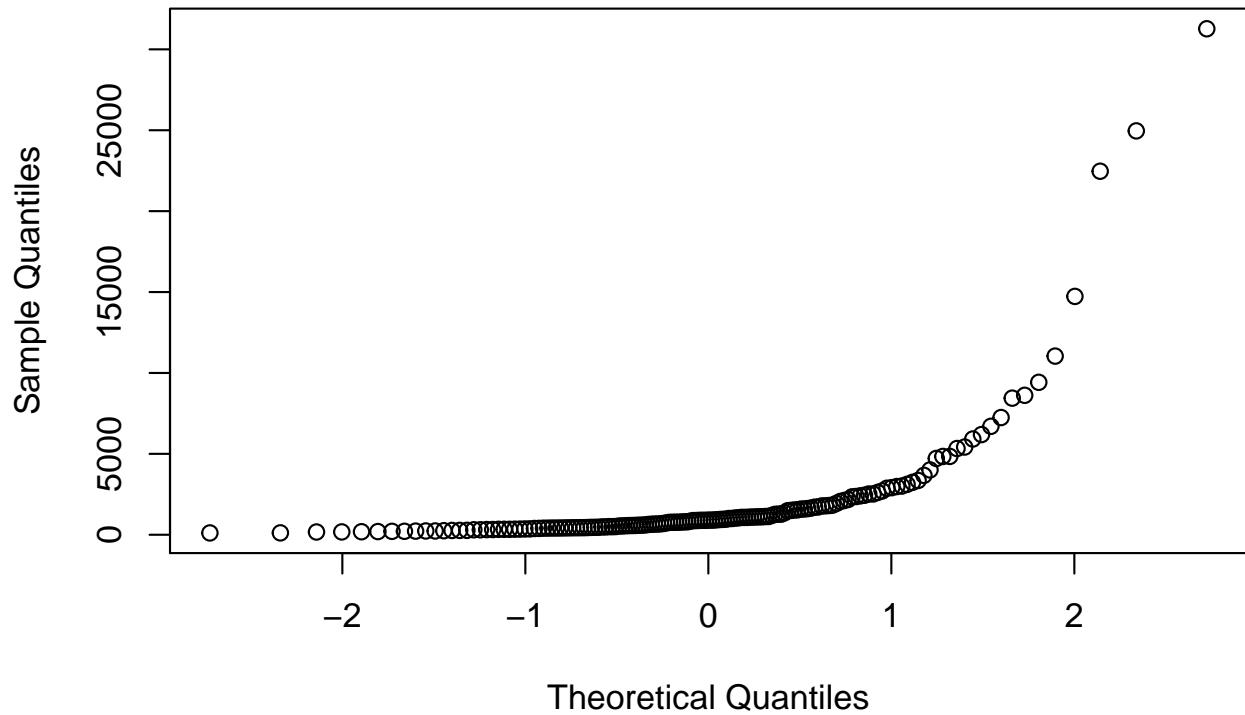
```
hist(log(Dificam))
```

## Histogram of log(Difacam)



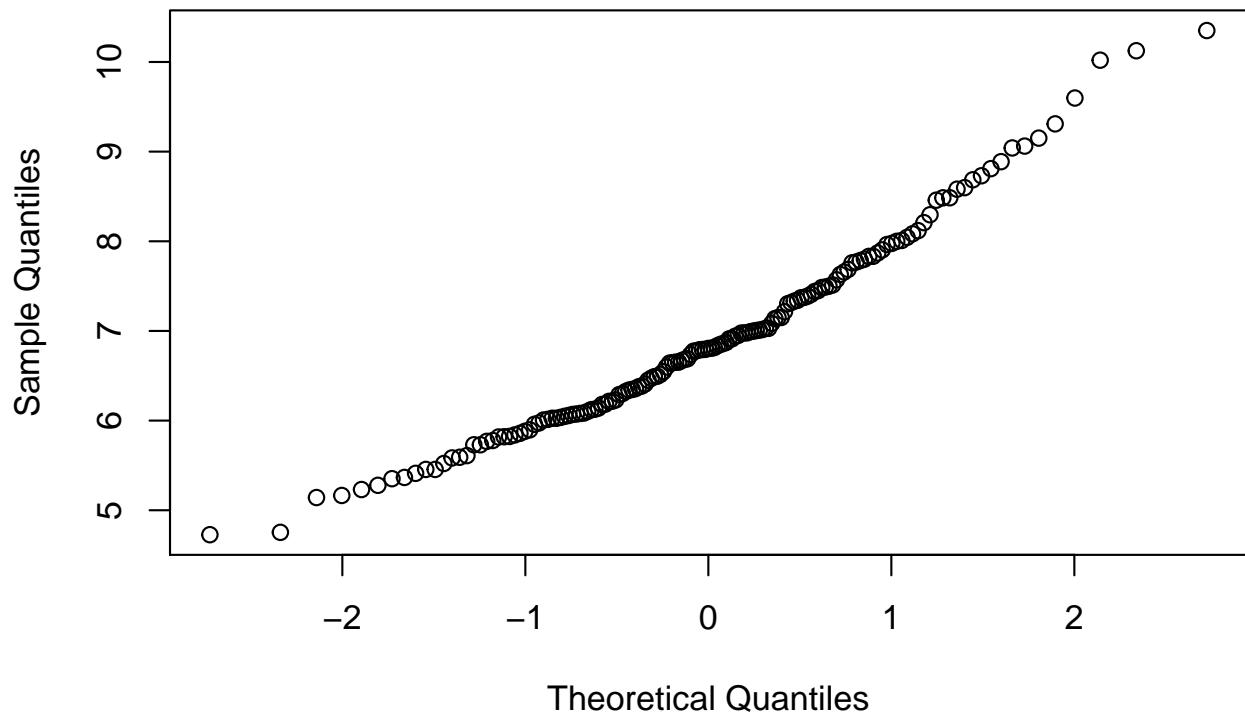
```
shapiro.test(Difacam)
##
## Shapiro-Wilk normality test
##
## data: Difacam
## W = 0.44601, p-value < 2.2e-16
shapiro.test(log(Difacam))
##
## Shapiro-Wilk normality test
##
## data: log(Difacam)
## W = 0.96842, p-value = 0.001267
qqnorm(Difacam)
```

### Normal Q–Q Plot



```
qqnorm(log(Difcam))
```

### Normal Q–Q Plot



```
Dificam_log <- log(Dificam)
censoT <- censoT %>% mutate("Dificam_log" = log(Dificam))
```

## 6.4 Análisis de Vecindad

```
# Vecindad por Contigüidad: Se crea objeto de vecindad con el criterio donde todos son vecinos
censoT.sp <- as_Spatial(censoT)
censoT.np <- poly2nb(censoT.sp, queen = TRUE)
summary(censoT.np)
## Neighbour list object:
## Number of regions: 155
## Number of nonzero links: 804
## Percentage nonzero weights: 3.346514
## Average number of links: 5.187097
## Link number distribution:
##
##   1   2   3   4   5   6   7   8   9  10  11  12  14
##  1 10 20 34 33 22 13 13  4   1   1   2   1
## 1 least connected region:
## 107 with 1 link
## 1 most connected region:
## 63 with 14 links

# Evaluación de cardinalidad
card(censoT.np)
## [1] 4 6 3 6 7 6 3 2 6 5 5 7 5 8 4 3 4 7 4 3 5 6 4
## [24] 4 4 8 5 5 4 6 4 5 12 4 5 6 7 4 5 3 3 4 4 6 5 8
## [47] 3 10 3 7 3 2 3 3 5 7 8 5 2 4 5 3 14 8 7 3 6 2 5
## [70] 4 5 4 8 5 2 3 5 2 6 3 8 7 4 5 6 4 5 4 2 7 4 4
## [93] 2 7 2 9 3 3 5 6 5 2 6 11 4 8 1 7 5 4 6 5 5 5 4
## [116] 9 6 7 4 12 5 5 4 8 4 3 5 4 9 4 3 8 6 5 9 4 6 8
## [139] 7 6 8 4 6 6 4 8 3 6 5 4 6 4 5 5 5
sapply(censoT.np, function(x) x)
## [[1]]
## [1] 149 150 151 154
##
## [[2]]
## [1] 6 7 8 9 11 27
##
## [[3]]
## [1] 11 79 146
##
## [[4]]
## [1] 5 6 9 14 21 104
##
```

```

## [[5]]
## [1] 4 6 10 64 65 104 105
##
## [[6]]
## [1] 2 4 5 9 10 11
##
## [[7]]
## [1] 2 8 9
##
## [[8]]
## [1] 2 7
##
## [[9]]
## [1] 2 4 6 7 21 27
##
## [[10]]
## [1] 5 6 11 64 146
##
## [[11]]
## [1] 2 3 6 10 146
##
## [[12]]
## [1] 13 14 15 56 57 106 109
##
## [[13]]
## [1] 12 14 56 57 109
##
## [[14]]
## [1] 4 12 13 21 22 56 104 109
##
## [[15]]
## [1] 12 16 57 106
##
## [[16]]
## [1] 15 55 106
##
## [[17]]
## [1] 18 23 24 26
##
## [[18]]
## [1] 17 22 23 24 25 26 56
##
## [[19]]
## [1] 20 26 77 78
##
## [[20]]
## [1] 19 23 26

```

```

## 
## [[21]]
## [1] 4 9 14 22 27
##
## [[22]]
## [1] 14 18 21 24 27 56
##
## [[23]]
## [1] 17 18 20 26
##
## [[24]]
## [1] 17 18 22 27
##
## [[25]]
## [1] 18 26 56 57
##
## [[26]]
## [1] 17 18 19 20 23 25 57 77
##
## [[27]]
## [1] 2 9 21 22 24
##
## [[28]]
## [1] 29 30 71 74 75
##
## [[29]]
## [1] 28 30 31 32
##
## [[30]]
## [1] 28 29 32 73 74 129
##
## [[31]]
## [1] 29 32 44 130
##
## [[32]]
## [1] 29 30 31 129 130
##
## [[33]]
## [1] 35 36 38 50 63 67 69 70 90 91 92 118
##
## [[34]]
## [1] 37 67 69 94
##
## [[35]]
## [1] 33 36 37 39 69
##
## [[36]]

```

```

## [1] 33 35 38 39 116 119
##
## [[37]]
## [1] 34 35 39 69 94 117 140
##
## [[38]]
## [1] 33 36 118 119
##
## [[39]]
## [1] 35 36 37 116 117
##
## [[40]]
## [1] 41 42 108
##
## [[41]]
## [1] 40 44 108
##
## [[42]]
## [1] 40 43 45 108
##
## [[43]]
## [1] 42 45 54 55
##
## [[44]]
## [1] 31 41 104 108 129 130
##
## [[45]]
## [1] 42 43 55 106 108
##
## [[46]]
## [1] 47 58 61 112 113 143 144 145
##
## [[47]]
## [1] 46 58 144
##
## [[48]]
## [1] 49 50 51 63 81 87 90 123 125 127
##
## [[49]]
## [1] 48 63 90
##
## [[50]]
## [1] 33 48 51 70 87 90 91
##
## [[51]]
## [1] 48 50 87
##

```

```

## [[52]]
## [1] 53 54
##
## [[53]]
## [1] 52 57 77
##
## [[54]]
## [1] 43 52 55
##
## [[55]]
## [1] 16 43 45 54 106
##
## [[56]]
## [1] 12 13 14 18 22 25 57
##
## [[57]]
## [1] 12 13 15 25 26 53 56 77
##
## [[58]]
## [1] 46 47 59 60 61
##
## [[59]]
## [1] 58 60
##
## [[60]]
## [1] 58 59 61 62
##
## [[61]]
## [1] 46 58 60 62 112
##
## [[62]]
## [1] 60 61 112
##
## [[63]]
## [1] 33 48 49 64 65 66 90 92 118 120 122 127 128 135
##
## [[64]]
## [1] 5 10 63 65 135 146 147 148
##
## [[65]]
## [1] 5 63 64 105 122 124 135
##
## [[66]]
## [1] 63 118 135
##
## [[67]]
## [1] 33 34 68 69 70 94

```

```

## 
## [[68]]
## [1] 67 70
##
## [[69]]
## [1] 33 34 35 37 67
##
## [[70]]
## [1] 33 50 67 68
##
## [[71]]
## [1] 28 72 74 75 76
##
## [[72]]
## [1] 71 73 74 76
##
## [[73]]
## [1] 30 72 74 76 88 129 132 134
##
## [[74]]
## [1] 28 30 71 72 73
##
## [[75]]
## [1] 28 71
##
## [[76]]
## [1] 71 72 73
##
## [[77]]
## [1] 19 26 53 57 78
##
## [[78]]
## [1] 19 77
##
## [[79]]
## [1] 3 80 99 101 103 146
##
## [[80]]
## [1] 79 97 101
##
## [[81]]
## [1] 48 82 84 86 87 89 120 125
##
## [[82]]
## [1] 81 83 84 120 121 126 133
##
## [[83]]

```

```

## [1] 82 84 85 133
##
## [[84]]
## [1] 81 82 83 85 86
##
## [[85]]
## [1] 83 84 86 88 133 134
##
## [[86]]
## [1] 81 84 85 88
##
## [[87]]
## [1] 48 50 51 81 89
##
## [[88]]
## [1] 73 85 86 134
##
## [[89]]
## [1] 81 87
##
## [[90]]
## [1] 33 48 49 50 63 91 92
##
## [[91]]
## [1] 33 50 90 92
##
## [[92]]
## [1] 33 63 90 91
##
## [[93]]
## [1] 94 95
##
## [[94]]
## [1] 34 37 67 93 95 140 144
##
## [[95]]
## [1] 93 94
##
## [[96]]
## [1] 97 98 99 100 101 102 150 154 155
##
## [[97]]
## [1] 80 96 101
##
## [[98]]
## [1] 96 102 150
##

```

```

## [[99]]
## [1] 79 96 100 101 103
##
## [[100]]
## [1] 96 99 103 137 141 155
##
## [[101]]
## [1] 79 80 96 97 99
##
## [[102]]
## [1] 96 98
##
## [[103]]
## [1] 79 99 100 137 146 148
##
## [[104]]
## [1] 4 5 14 44 105 106 107 108 109 124 129
##
## [[105]]
## [1] 5 65 104 124
##
## [[106]]
## [1] 12 15 16 45 55 104 108 109
##
## [[107]]
## [1] 104
##
## [[108]]
## [1] 40 41 42 44 45 104 106
##
## [[109]]
## [1] 12 13 14 104 106
##
## [[110]]
## [1] 112 113 114 115
##
## [[111]]
## [1] 114 115 139 143 152 153
##
## [[112]]
## [1] 46 61 62 110 113
##
## [[113]]
## [1] 46 110 112 114 143
##
## [[114]]
## [1] 110 111 113 115 143

```

```

## 
## [[115]]
## [1] 110 111 114 152
##
## [[116]]
## [1] 36 39 117 118 119 135 136 141 142
##
## [[117]]
## [1] 37 39 116 138 140 142
##
## [[118]]
## [1] 33 38 63 66 116 119 135
##
## [[119]]
## [1] 36 38 116 118
##
## [[120]]
## [1] 63 81 82 121 122 123 124 125 126 127 128 132
##
## [[121]]
## [1] 82 120 126 132 133
##
## [[122]]
## [1] 63 65 120 124 128
##
## [[123]]
## [1] 48 120 125 127
##
## [[124]]
## [1] 65 104 105 120 122 129 131 132
##
## [[125]]
## [1] 48 81 120 123
##
## [[126]]
## [1] 82 120 121
##
## [[127]]
## [1] 48 63 120 123 128
##
## [[128]]
## [1] 63 120 122 127
##
## [[129]]
## [1] 30 32 44 73 104 124 130 131 132
##
## [[130]]

```

```

## [1] 31 32 44 129
##
## [[131]]
## [1] 124 129 132
##
## [[132]]
## [1] 73 120 121 124 129 131 133 134
##
## [[133]]
## [1] 82 83 85 121 132 134
##
## [[134]]
## [1] 73 85 88 132 133
##
## [[135]]
## [1] 63 64 65 66 116 118 136 137 148
##
## [[136]]
## [1] 116 135 137 141
##
## [[137]]
## [1] 100 103 135 136 141 148
##
## [[138]]
## [1] 117 139 140 141 142 149 151 153
##
## [[139]]
## [1] 111 138 140 143 144 145 153
##
## [[140]]
## [1] 37 94 117 138 139 144
##
## [[141]]
## [1] 100 116 136 137 138 142 151 155
##
## [[142]]
## [1] 116 117 138 141
##
## [[143]]
## [1] 46 111 113 114 139 145
##
## [[144]]
## [1] 46 47 94 139 140 145
##
## [[145]]
## [1] 46 139 143 144
##

```

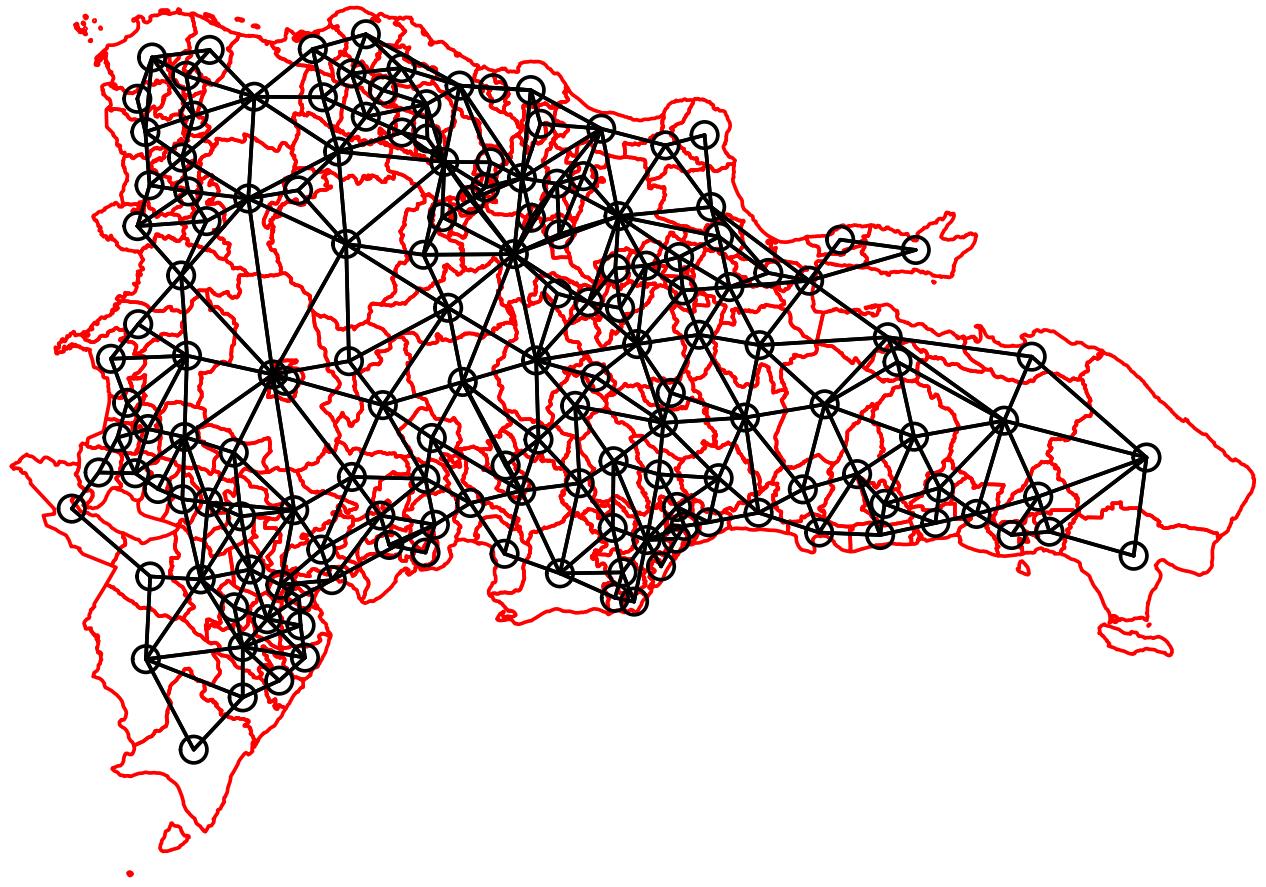
```

## [[146]]
## [1] 3 10 11 64 79 103 147 148
##
## [[147]]
## [1] 64 146 148
##
## [[148]]
## [1] 64 103 135 137 146 147
##
## [[149]]
## [1] 1 138 151 152 153
##
## [[150]]
## [1] 1 96 98 154
##
## [[151]]
## [1] 1 138 141 149 154 155
##
## [[152]]
## [1] 111 115 149 153
##
## [[153]]
## [1] 111 138 139 149 152
##
## [[154]]
## [1] 1 96 150 151 155
##
## [[155]]
## [1] 96 100 141 151 154

#resultado: Es simetrico
is.symmetric.nb(censoT.np)
## [1] TRUE

plot(censoT.sp, border="red", lwd=1)
plot(censoT.np, coordinates(censoT.sp), add=T)

```



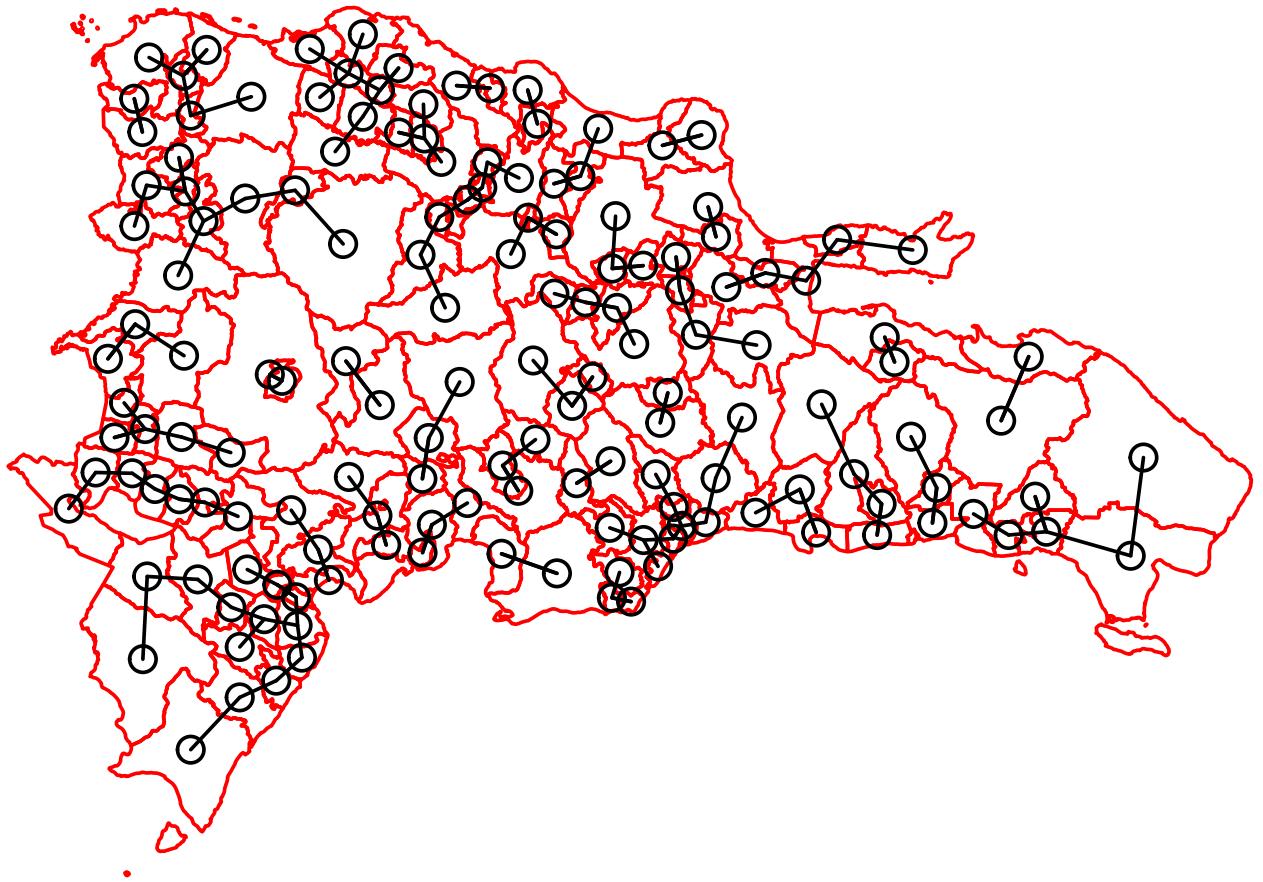
```

# Vecindad por numero de vecinos: Se crea objeto de vecindad donde cada municipio tenga solo un
coords <- coordinates(censoT.sp)
ident <- row.names(censoT.sp)
censoT.np.k1 <- knn2nb(knearneigh(coords,k=1), row.names = ident)
censoT.np.k2 <- knn2nb(knearneigh(coords,k=2), row.names = ident)

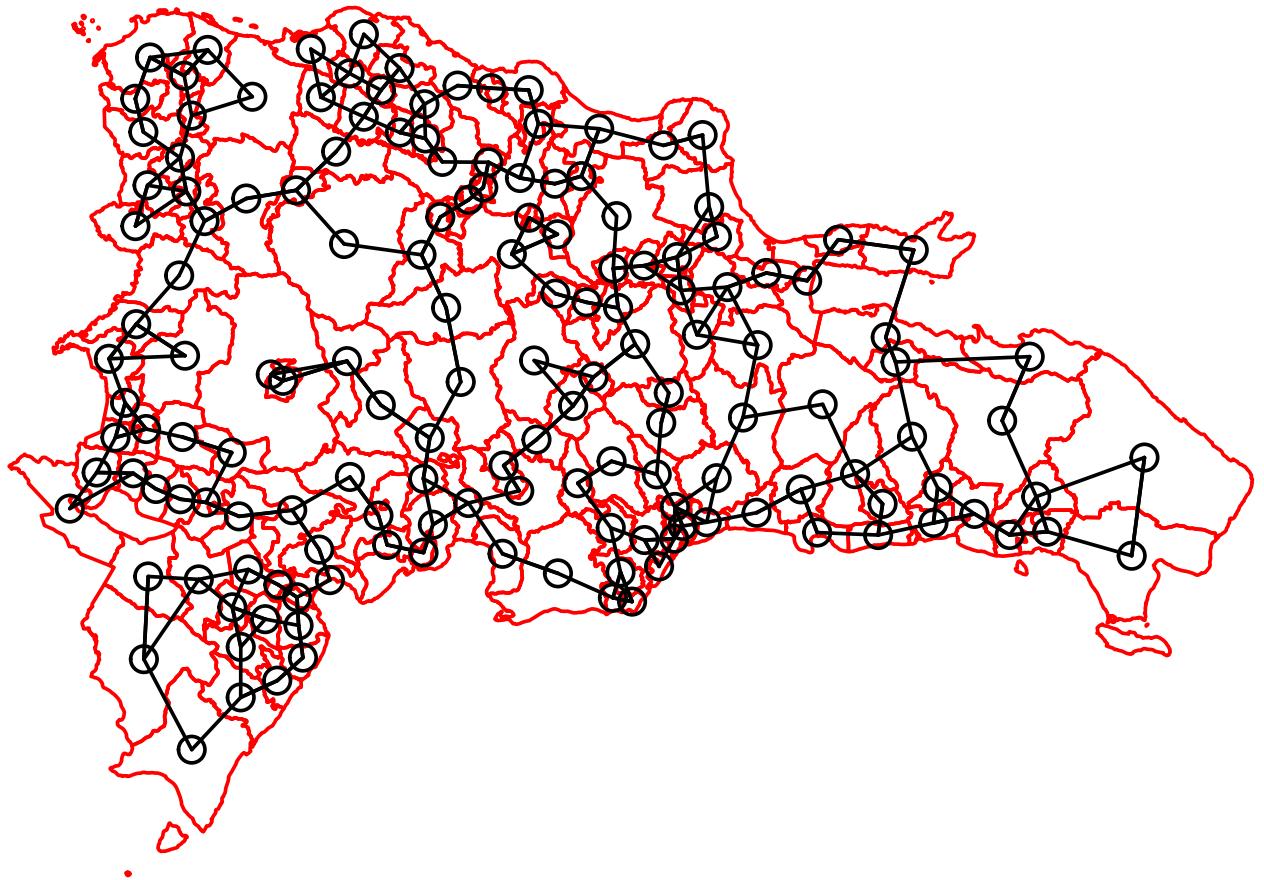
is.symmetric.nb(censoT.np.k1)
## [1] FALSE
is.symmetric.nb(censoT.np.k2)
## [1] FALSE

# Resultado: En ambos casos NO es simetrico
plot(censoT.sp, border="red", lwd=1)
plot(censoT.np.k1, coordinates(censoT.sp), add=T)

```

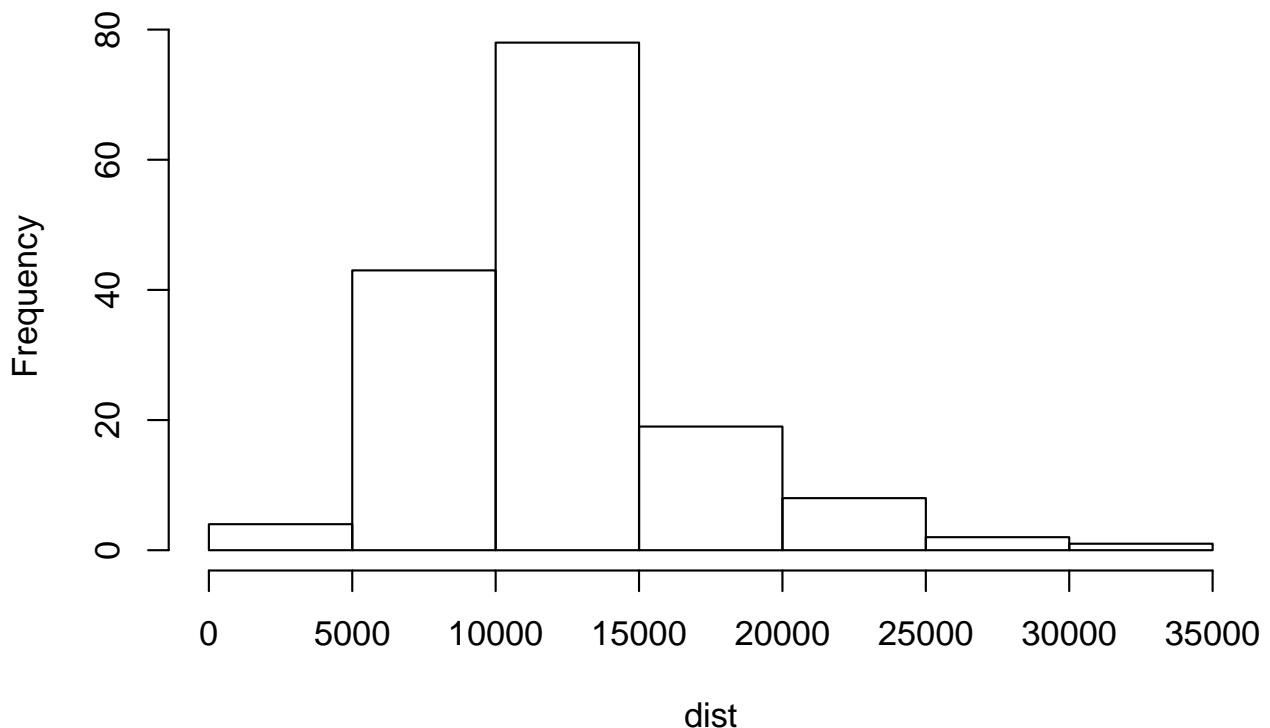


```
plot(censoT.sp, border="red", lwd=1)
plot(censoT.np.k2, coordinates(censoT.sp), add=T)
```

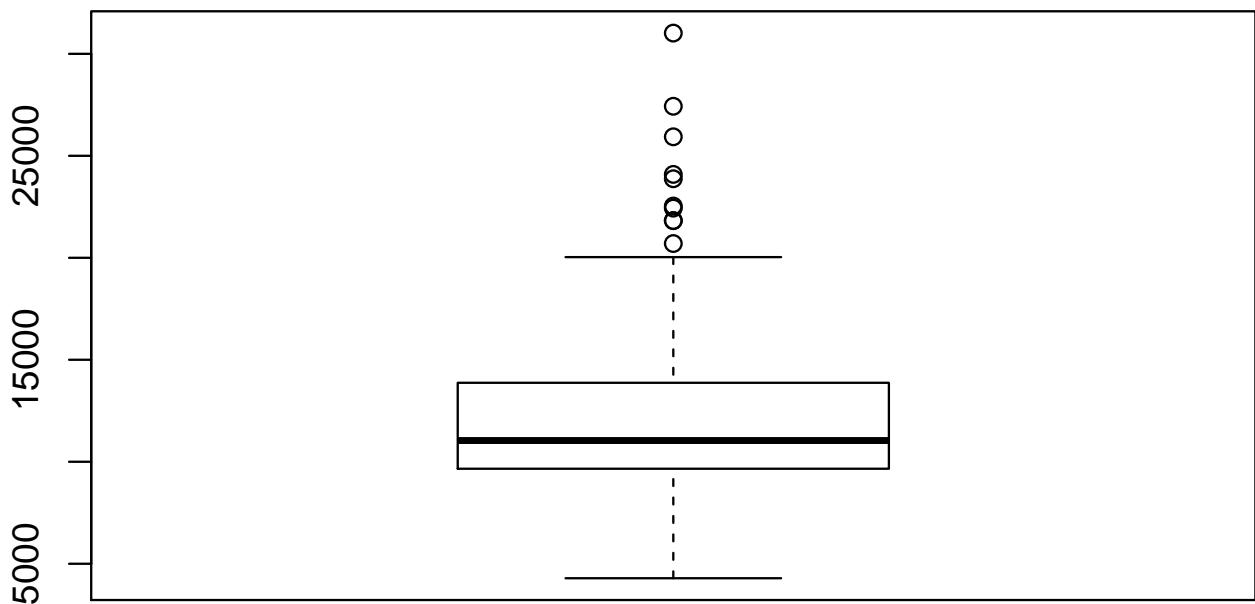


```
# Determinamos la distancia máxima y mínima al vecino mas proximo usando k=1
dist <- unlist(nbdists(censoT.np.k1, coords))
summary(dist)
##      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
##     4290    9657   11042   12320   13872   31017
hist(dist)
```

### Histogram of dist



```
boxplot(dist)
```



```
rownames(censoT) <- censoT$TOPONIMIA  
nb <- poly2nb(censoT)  
summary(nb)  
## Neighbour list object:  
## Number of regions: 155
```

```

## Number of nonzero links: 804
## Percentage nonzero weights: 3.346514
## Average number of links: 5.187097
## Link number distribution:
##
##   1   2   3   4   5   6   7   8   9  10  11  12  14
## 1 10 20 34 33 22 13 13  4  1  1  2  1
## 1 least connected region:
## JUAN DE HERRERA with 1 link
## 1 most connected region:
## LA VEGA with 14 links

```

## 6.5 Ponderadores espaciales (*Construcción de objeto de pesos*)

```

# Estilo Weighted. Donde los pesos de las observaciones vecinas suman 1. Estandarización por fila
censo.w.w <- nb2listw(nb)
censo.w.w
## Characteristics of weights list object:
## Neighbour list object:
## Number of regions: 155
## Number of nonzero links: 804
## Percentage nonzero weights: 3.346514
## Average number of links: 5.187097
##
## Weights style: W
## Weights constants summary:
##      n      nn     S0      S1      S2
## W 155 24025 155 65.94606 650.7687

# Estilo Binario. Donde los pesos son indicativos de la relación entre dos o mas observaciones.
censo.w.b <- nb2listw(nb, style = 'B')
censo.w.b
## Characteristics of weights list object:
## Neighbour list object:
## Number of regions: 155
## Number of nonzero links: 804
## Percentage nonzero weights: 3.346514
## Average number of links: 5.187097
##
## Weights style: B
## Weights constants summary:
##      n      nn     S0      S1      S2
## B 155 24025 804 1608 19520

```

## 6.6 Autocorrelacion Espacial

```
# Prueba de Breuch-Pagan
coordsxy <- censoT %>%
  st_centroid() %>%
  mutate (x=unlist(map(geom,1)),
         y=unlist(map(geom,2))) %>%
  st_drop_geometry() %>%
  select(ENLACE, x, y)
## Warning in st_centroid.sf(.): st_centroid assumes attributes are constant
## over geometries of x
coordsxy
## # ENLACE      x      y
## 1 100101 400576.8 2044091
## 2 050201 308829.1 2037708
## 3 050202 336675.2 2034232
## 4 050203 289007.6 2058192
## 5 050204 298570.6 2080773
## 6 050205 311899.8 2057722
## 7 050206 300563.9 2037000
## 8 050207 311931.4 2034599
## 9 050208 297803.0 2046138
## 10 050209 313902.6 2070484
## 11 050210 325993.3 2050137
## 12 060301 248793.8 2047108
## 13 060302 254183.2 2045951
## 14 060303 270830.3 2047876
## 15 060304 235833.2 2051310
## 16 060305 228193.1 2054516
## 17 060401 272818.5 2011818
## 18 060402 262391.6 2013756
## 19 060403 254722.0 1989280
## 20 060404 266143.2 1994524
## 21 060405 279288.5 2035539
## 22 060406 266486.1 2024529
## 23 060407 274187.1 2001728
## 24 060408 272335.4 2020630
## 25 060409 252085.2 2017487
## 26 060410 254728.5 2005072
## 27 060411 282623.2 2025979
## 28 040501 224236.2 2166258
## 29 040502 225565.7 2149580
## 30 040503 235714.0 2158230
## 31 040504 221765.7 2136807
## 32 040505 237696.6 2147659
## 33 030601 372356.7 2139919
## 34 030602 419244.2 2122143
## 35 030603 391313.6 2127124
```

```
## 36 030604 380987.7 2124416
## 37 030605 407069.2 2117718
## 38 030606 371332.0 2123535
## 39 030607 392510.6 2116668
## 40 070701 213454.7 2095259
## 41 070702 222013.0 2106096
## 42 070703 218529.2 2081475
## 43 070704 215517.5 2070627
## 44 070705 235443.1 2121345
## 45 070706 225100.7 2073278
## 46 080801 493062.0 2075900
## 47 080802 501691.0 2095945
## 48 010901 342119.5 2151861
## 49 010902 344981.4 2139419
## 50 010903 366931.6 2167290
## 51 010904 347911.7 2168817
## 52 061001 201206.0 2048370
## 53 061002 225759.6 2027135
## 54 061003 209629.1 2059673
## 55 061004 221071.3 2059431
## 56 061005 257075.9 2029308
## 57 061006 241582.7 2026219
## 58 081101 537616.5 2064370
## 59 081102 533606.2 2033611
## 60 081201 517505.0 2026984
## 61 081202 503716.1 2052116
## 62 081203 495423.5 2040118
## 63 021301 339579.4 2128029
## 64 021302 323578.8 2088026
## 65 021303 319010.5 2111237
## 66 021304 353216.5 2115675
## 67 031401 401356.9 2142772
## 68 031402 399289.8 2165300
## 69 031403 403801.2 2133465
## 70 031404 387038.9 2162067
## 71 041501 226193.1 2189714
## 72 041502 237018.4 2183857
## 73 041503 258355.1 2177340
## 74 041504 239378.2 2171396
## 75 041505 221719.9 2176656
## 76 041506 244390.5 2192066
## 77 061601 224739.2 1998913
## 78 061602 239376.1 1972918
## 79 051701 354008.1 2028099
## 80 051702 371207.6 2020451
## 81 011801 322604.0 2180778
## 82 011802 312232.7 2174944
```

```
## 83 011803 298607.2 2179332
## 84 011804 304447.3 2186274
## 85 011805 288791.7 2184581
## 86 011806 293257.9 2196858
## 87 011807 344859.0 2179378
## 88 011808 276658.7 2192229
## 89 011809 333073.6 2179974
## 90 031901 353080.8 2150071
## 91 031902 361367.4 2152437
## 92 031903 353834.8 2134247
## 93 032001 465392.7 2129343
## 94 032002 432022.2 2119706
## 95 032003 441731.3 2132547
## 96 052101 381121.1 2038534
## 97 052102 377220.0 2019264
## 98 052103 390429.0 2038970
## 99 052104 370532.9 2042348
## 100 052105 370802.0 2063198
## 101 052106 373604.5 2028413
## 102 052107 385655.3 2030272
## 103 052108 360193.7 2056318
## 104 072201 263934.3 2090465
## 105 072202 287975.4 2094527
## 106 072203 236485.3 2070777
## 107 072204 267940.0 2088369
## 108 072205 237275.6 2096238
## 109 072206 251918.6 2065917
## 110 092301 471552.1 2043872
## 111 092302 447037.2 2059167
## 112 092303 484408.5 2046716
## 113 092304 472875.4 2054872
## 114 092305 455634.6 2049854
## 115 092306 454233.7 2040098
## 116 022401 378201.0 2099880
## 117 022402 397518.5 2102781
## 118 022403 362986.1 2112936
## 119 022404 372846.6 2111201
## 120 012501 318262.6 2155242
## 121 012502 304538.5 2166095
## 122 012503 311309.9 2127829
## 123 012504 330650.3 2148915
## 124 012505 287084.6 2131214
## 125 012506 332234.0 2156780
## 126 012507 312521.8 2164237
## 127 012508 326060.0 2145023
## 128 012509 317192.6 2139618
## 129 042601 256458.7 2145440
```

```

## 130 042602 243396.5 2138360
## 131 042603 272050.7 2148004
## 132 042701 284585.6 2160123
## 133 042702 293252.9 2171034
## 134 042703 279807.8 2177125
## 135 022801 346550.6 2094746
## 136 022802 365140.2 2089549
## 137 022803 358761.9 2080578
## 138 092901 411938.7 2076885
## 139 092902 436884.1 2081013
## 140 092903 416514.7 2099541
## 141 092904 386347.8 2075230
## 142 092905 388715.6 2084593
## 143 093001 464859.2 2070897
## 144 093002 456632.7 2102023
## 145 093003 459686.8 2094264
## 146 053101 341909.1 2053945
## 147 053102 337015.2 2061850
## 148 053103 347299.1 2069964
## 149 103201 416195.1 2047076
## 150 103202 393011.0 2042924
## 151 103203 403690.6 2057900
## 152 103204 435165.8 2040880
## 153 103205 430038.2 2054384
## 154 103206 390626.7 2048896
## 155 103207 384940.5 2059089
censoT <- censoT %>% inner_join(coordsxy)
## Joining, by = "ENLACE"

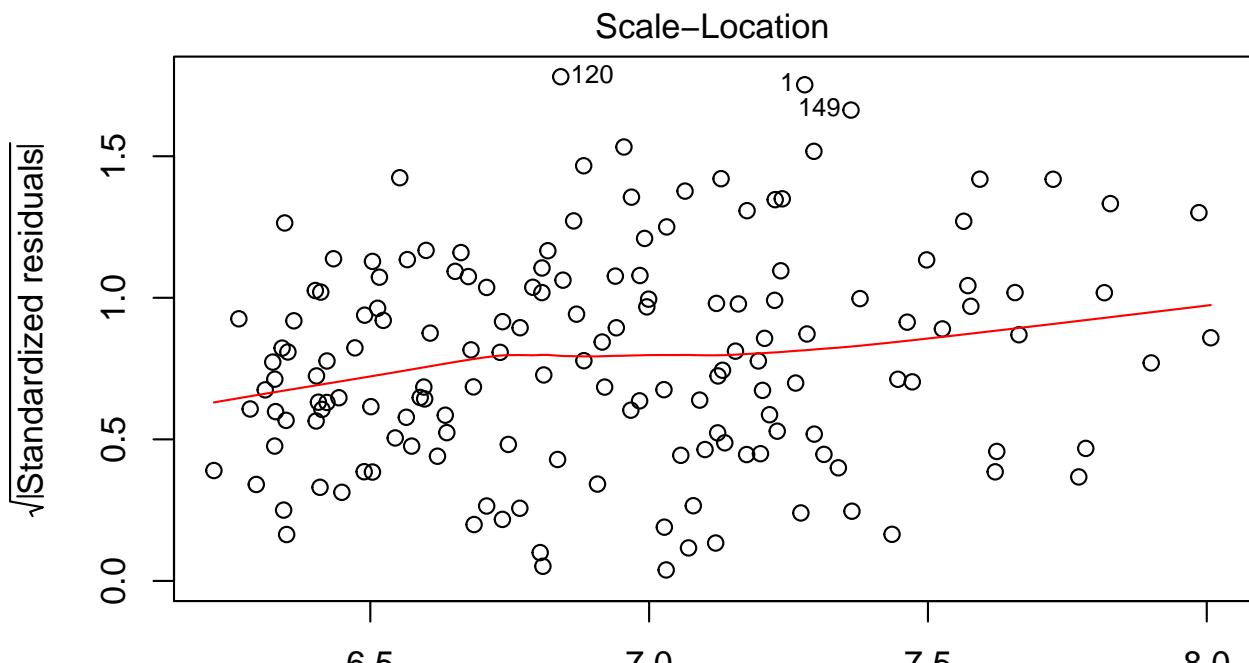
```

## 6.7 Prueba de homocedasticidad de la variable transformada

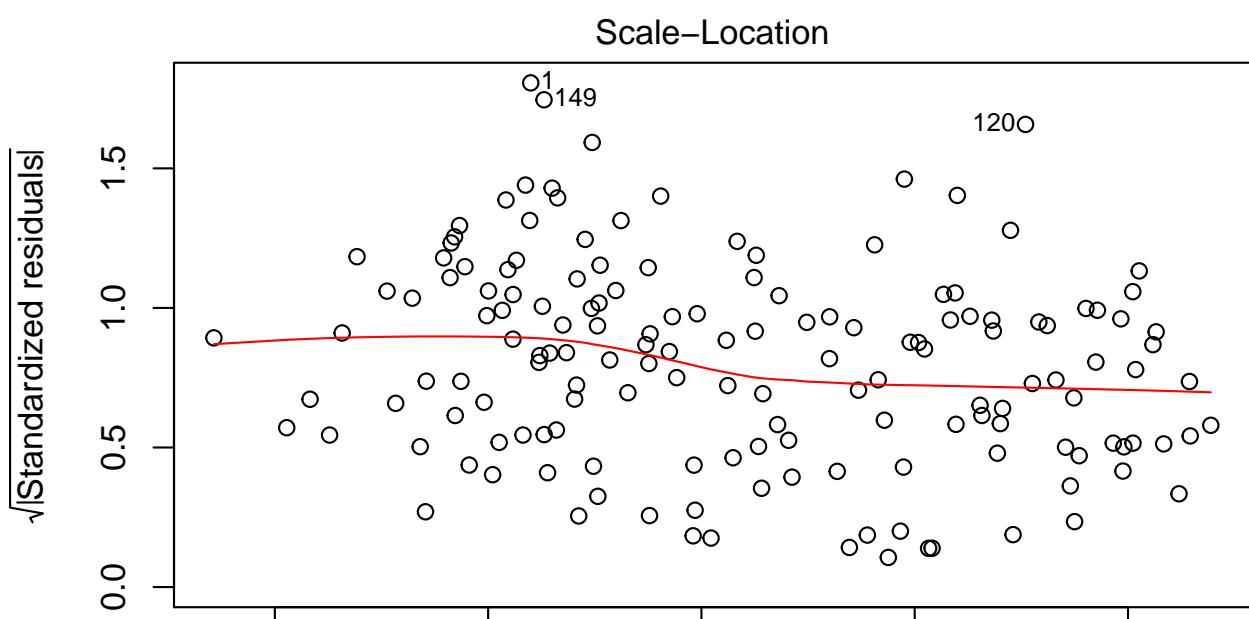
```

censoT %>% lm(Difcam_log~ x, .) %>% plot(3)

```



```
censoT %>% lm(Difacam_log~ y, .) %>% plot(3)
```



```

censoT %>% lm(Difcam_log ~ x, .) %>% bptest() #Es mas homocedastica
##
## studentized Breusch-Pagan test
##
## data: .
## BP = 5.6667, df = 1, p-value = 0.01729
censoT %>% lm(Difcam_log ~ y, .) %>% bptest()
##
## studentized Breusch-Pagan test
##
## data: .
## BP = 4.6633, df = 1, p-value = 0.03081
#El indice de significancia es menor de 0.05, por tanto se rechaza la hipotesis nula

```

## 6.8 Test de I de Moran global

```

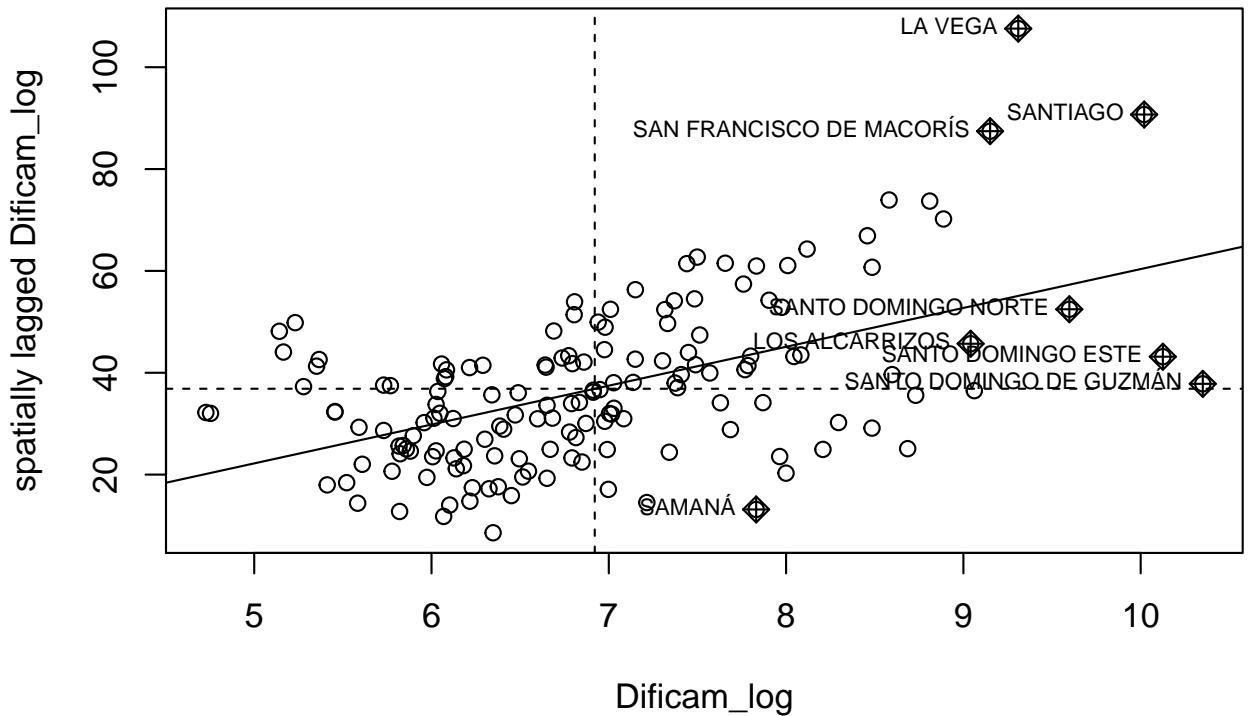
moran.test(x=Difcam_log, listw = censo.w.w, na.action = na.omit)
##
## Moran I test under randomisation
##
## data: Difcam_log
## weights: censo.w.w
##
## Moran I statistic standard deviate = 6.615, p-value = 1.858e-11
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic      Expectation      Variance
##          0.333709240     -0.006493506     0.002644972
moran.test(x=Difcam_log, listw = censo.w.b, na.action = na.omit) #Este nos da el valor mas cercano
##
## Moran I test under randomisation
##
## data: Difcam_log
## weights: censo.w.b
##
## Moran I statistic standard deviate = 8.0524, p-value = 4.061e-16
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic      Expectation      Variance
##          0.385389193     -0.006493506     0.002368463

# Para los pesos tanto estandarizado como binario, se comprueba que si hay correlacion espacial,

```

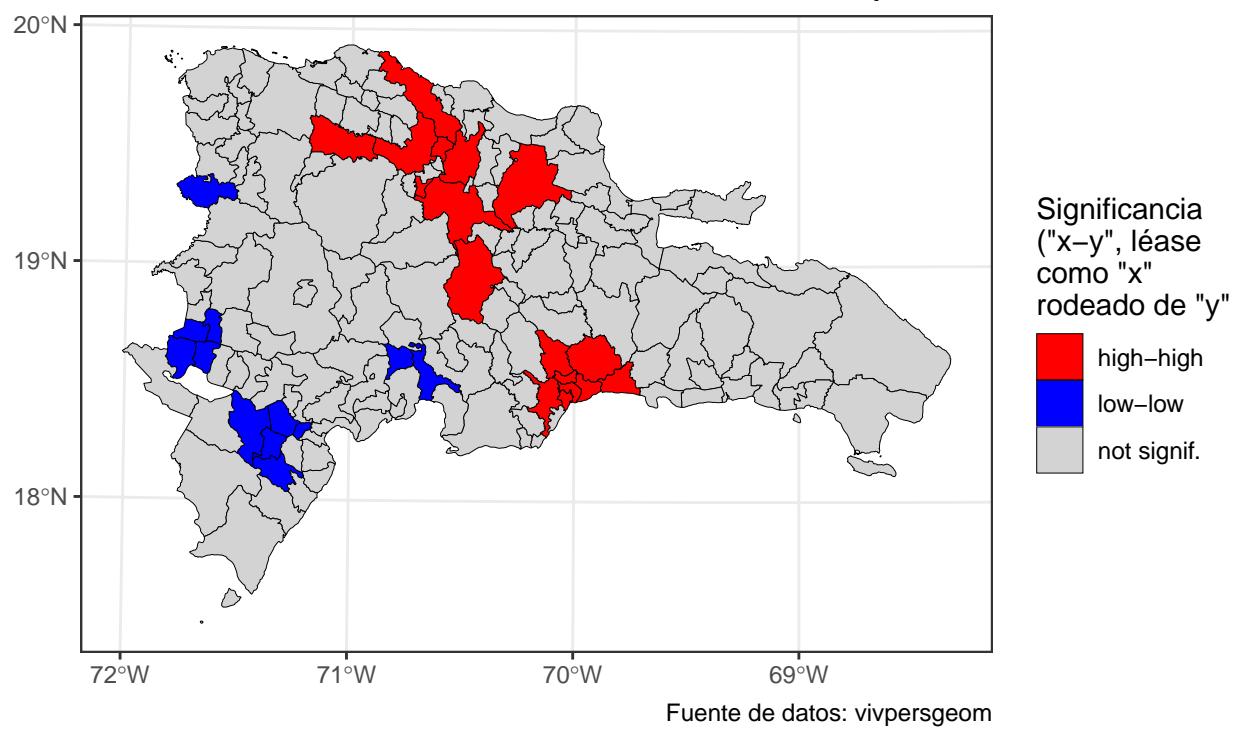
## 6.9 Test de I de Moran Local

```
moran.plot(x=Dificam_log, listw = censo.w.b)
```



```
DificamLoc <- localmoran(Dificam_log, listw = censo.w.b)
source('data/lisaclusters.R')
lisamap(objesp = censoT,
        var = 'Dificam_log',
        pesos = censo.w.b,
        tituloleyenda = 'Significancia\n("x-y", léase\ncomo "x"\nrodeado de "y")',
        leyenda = T,
        anchuratitulo = 1000,
        tamanotitulo = 16,
        fuentedatos = 'vivpersgeom',
        titulomap = paste0('Clusters LISA de Personas con dificultad para caminar o subir escal'),
        )
## $grafico
```

## Clusters LISA de Personas con dificultad para caminar o subir



```
##  
## $objeto  
## Simple feature collection with 155 features and 20 fields  
## geometry type:  MULTIPOLYGON  
## dimension:      XY  
## bbox:            xmin: 182215.8 ymin: 1933532 xmax: 571365.3 ymax: 2205216  
## epsg (SRID):   32619  
## proj4string:    +proj=utm +zone=19 +datum=WGS84 +units=m +no_defs  
## First 10 features:  
##           TOPONIMIA ENLACE  
## 1  SANTO DOMINGO DE GUZMÁN 100101  
## 2                  AZUA 050201  
## 3          LAS CHARCAS 050202  
## 4  LAS YAYAS DE VIAJAMA 050203  
## 5      PADRE LAS CASAS 050204  
## 6             PERALTA 050205  
## 7        SABANA YEGUA 050206  
## 8        PUEBLO VIEJO 050207  
## 9       TÁBARA ARRIBA 050208  
## 10        GUAYABAL 050209  
## Condición de ocupación: Ocupada con personas presentes  
## 1                      288362  
## 2                      22797  
## 3                      3147  
## 4                      4807  
## 5                      5243
```

## 6		3144
## 7		4828
## 8		2689
## 9		4108
## 10		1386
## Condición de ocupación: Desocupada Población total		
## 1	42200	965040
## 2	1920	91345
## 3	944	11243
## 4	798	17620
## 5	1345	20041
## 6	446	15257
## 7	594	19020
## 8	208	11235
## 9	426	17647
## 10	464	5263
## Dificultad para Caminar o subir escalones: Si		
## 1		31271
## 2		2616
## 3		323
## 4		770
## 5		804
## 6		320
## 7		675
## 8		266
## 9		437
## 10		234
## Asiste o asistió a la escuela: Nunca asistió		
## 1		45138
## 2		12617
## 3		2329
## 4		3442
## 5		4336
## 6		3528
## 7		2813
## 8		1629
## 9		3860
## 10		909
## Nivel educativo más alto al que asistió: Preprimaria		
## 1		73199
## 2		7718
## 3		965
## 4		1400
## 5		2027
## 6		1360
## 7		1079
## 8		961
## 9		1344

```

## 10                                240
##   Nivel educativo más alto al que asistió: Primaria o básica
## 1                               290061
## 2                               37701
## 3                               5361
## 4                               8439
## 5                               8458
## 6                               5736
## 7                               9246
## 8                               5270
## 9                               7916
## 10                              2458
##   Nivel educativo más alto al que asistió: Secundaria o media
## 1                               248102
## 2                               19003
## 3                               1616
## 4                               2757
## 5                               3104
## 6                               2823
## 7                               3676
## 8                               2133
## 9                               2869
## 10                              924
##   Nivel educativo más alto al que asistió: Universitaria o superior
## 1                               257590
## 2                               8601
## 3                               204
## 4                               541
## 5                               1075
## 6                               1020
## 7                               1048
## 8                               484
## 9                               602
## 10                              479
##   Cantviv PorcPersD PorcPersD_log Dificam_log      x      y
## 1    330562  3.240384   1.1756918  10.350446 400576.8 2044091
## 2    24717   2.863868   1.0521731   7.869402 308829.1 2037708
## 3     4091   2.872899   1.0553215   5.777652 336675.2 2034232
## 4     5605   4.370034   1.4747708   6.646391 289007.6 2058192
## 5     6588   4.011776   1.3892340   6.689599 298570.6 2080773
## 6     3590   2.097398   0.7406975   5.768321 311899.8 2057722
## 7     5422   3.548896   1.2666365   6.514713 300563.9 2037000
## 8     2897   2.367601   0.8618773   5.583496 311931.4 2034599
## 9     4534   2.476342   0.9067823   6.079933 297803.0 2046138
## 10    1850   4.446133   1.4920348   5.455321 313902.6 2070484
##                               geom puntuacionz lagpuntuacionz quad_sig
## 1  MULTIPOLYGON (((405218.1 20...  3.1437363   9.29680778  high-high
## 2  MULTIPOLYGON (((319065.3 20...  0.8695441  -6.77884923 not signif.

```

```

## 3 MULTIPOLYGON (((341415.3 20... -1.0478095 -0.09622568 not signif.
## 4 MULTIPOLYGON (((304058.1 20... -0.2515008 -0.38419412 not signif.
## 5 MULTIPOLYGON (((312890.8 20... -0.2118945 -0.22410831 not signif.
## 6 MULTIPOLYGON (((317370.6 20... -1.0563629 -3.71835917 low-low
## 7 MULTIPOLYGON (((306745.8 20... -0.3722002 -1.12696486 not signif.
## 8 MULTIPOLYGON (((310447.9 20... -1.2257782 0.49734386 not signif.
## 9 MULTIPOLYGON (((306556.7 20... -0.7707308 -2.10198421 not signif.
## 10 MULTIPOLYGON (((322129.5 20... -1.3432670 -2.13317172 not signif.

```

## 6.10 Resultado Obtenido

Se obtuvo un mapa LISA - Local Indicator of Spatial Asociation -, que nos muestra los clusters de municipios donde existen la mayor y menor incidencia de casos de personas con dificultad para caminar o subir escaleras. Encontramos un hotspot en la zona noroeste (municipios en rojo), donde existe mayor cantidad de personas con dificultad para caminar o subir escalones; en la región suroeste y fronteriza es donde existe menos incidencia (municipios en azul); mientras en los demás municipios (en gris), no se evidencia la incidencia de casos.

## 7 FASE 2 - GEOESTADISTICA / ANALISIS PUNTUAL

### 7.1 Metodología

Para efectuar este análisis espacial, se cargará la base espacial de municipios como referencia y el archivo de datos de precipitación de diferentes años suministrado por la ONAMET. Luego de efectuar el análisis exploratorio de los datos espaciales (ESDA), obtener las informaciones estadísticas básicas, histogramas, y efectua pruebas (como la de Shapiro Wilk) para comprobar que los datos del año seleccionado tenga una distribución normal, se generan los variogramas modelo, para seleccionar el idoneo y así efectuar la interpolación (kriging) y establecer visualmente los municipios con mayor precipitación en el país.

### 7.2 Carga de archivos de datos de precipitaciones. El archivo de municipios esta ya en memoria

```

prec <- st_read('data/onamet_prec_anual_sf.gpkg')
## Reading layer 'onamet_prec_anual_sf' from data source '/home/lbine/unidad-0-asignacion-99-mi-
## Simple feature collection with 25 features and 37 fields
## geometry type:  POINT
## dimension:      XY
## bbox:            xmin: -71.7 ymin: 18.067 xmax: -68.367 ymax: 19.85
## epsg (SRID):    4326
## proj4string:    +proj=longlat +datum=WGS84 +no_defs
st_crs(prec)
## Coordinate Reference System:
##   EPSG: 4326
##   proj4string: "+proj=longlat +datum=WGS84 +no_defs"
crswgs84utm <- 32619
precUtm <- prec %>% st_transform(crs = crswgs84utm)

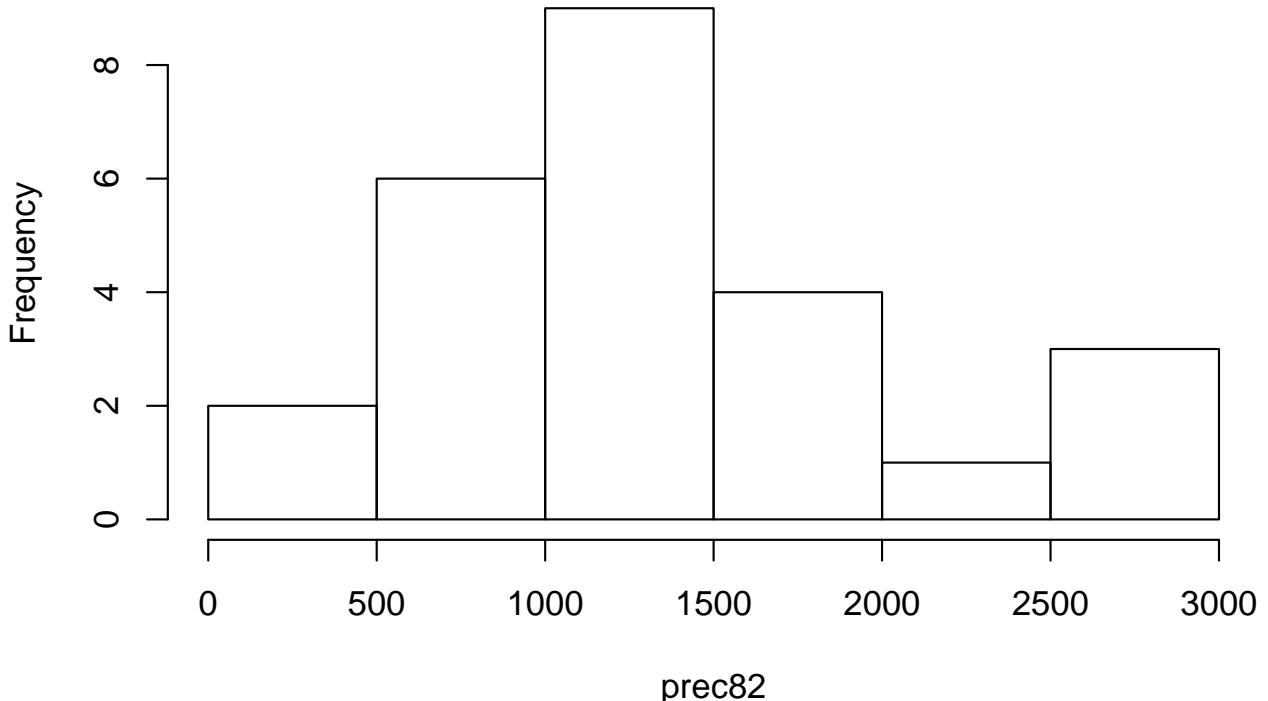
```

### 7.3 ESDA

```
nrow(precUtm)
## [1] 25
prec82 <- precUtm$a1982 # Selección de variable

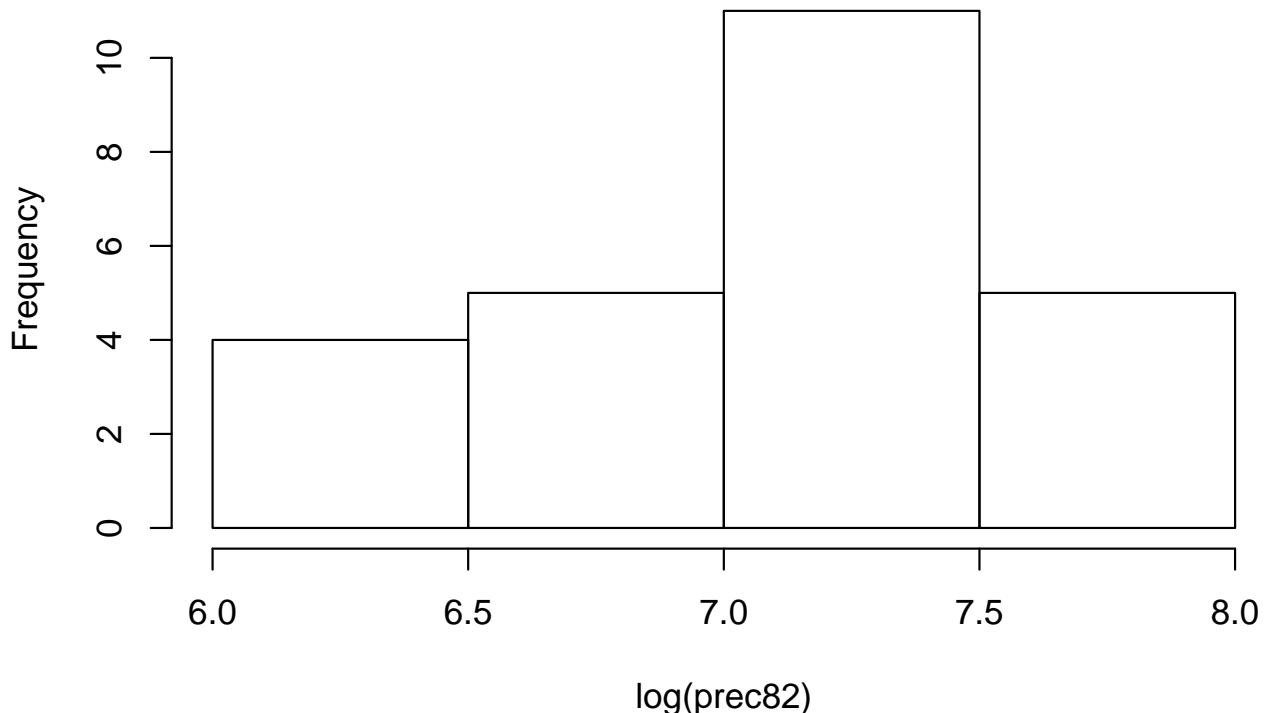
hist(prec82)
```

**Histogram of prec82**



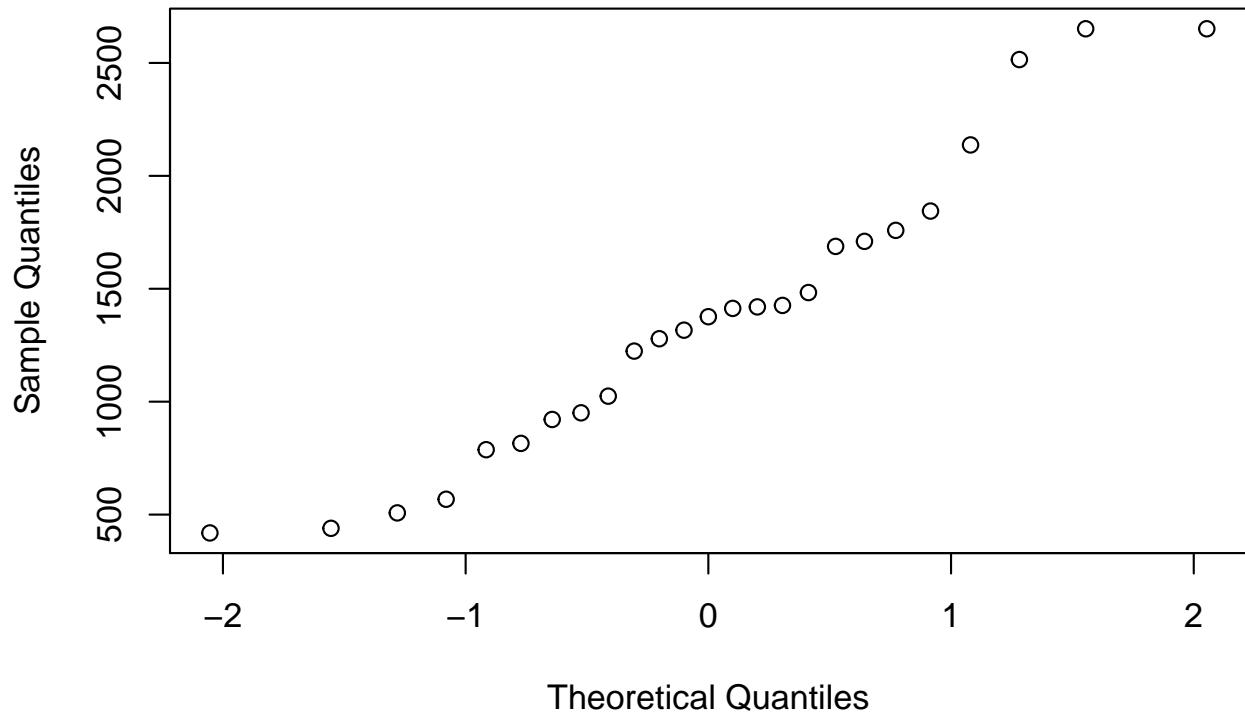
```
hist(log(prec82))
```

## Histogram of log(prec82)



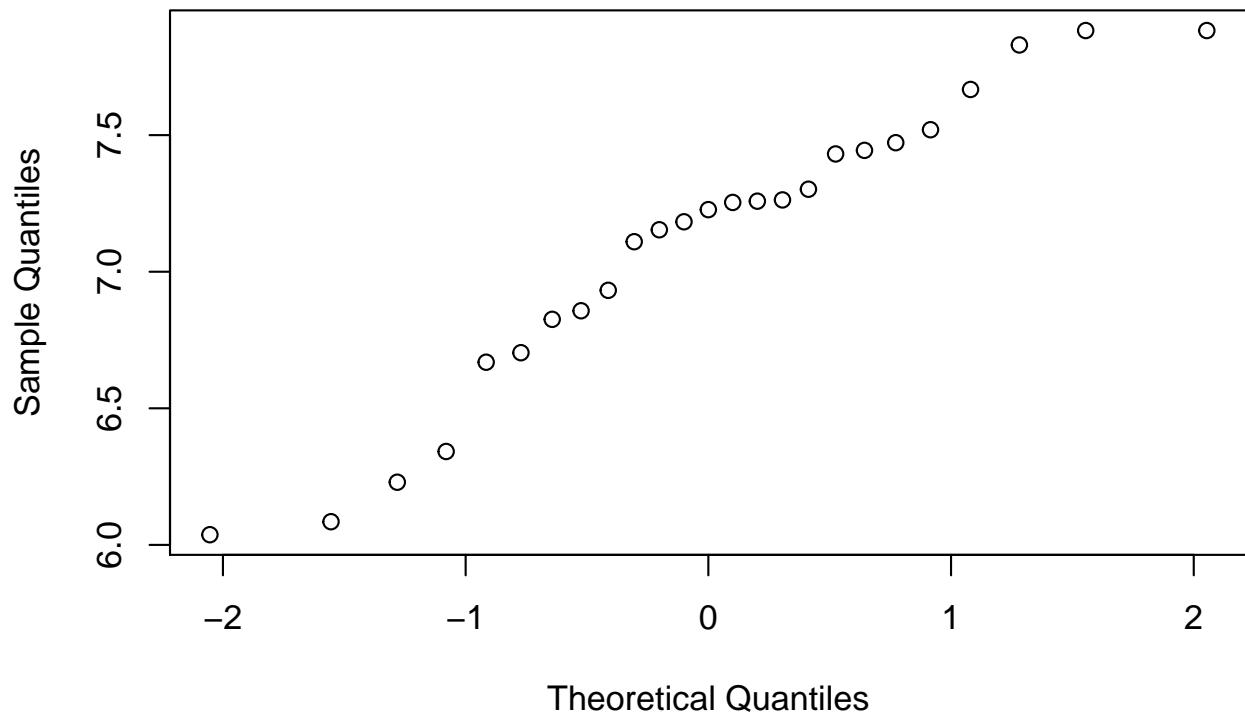
```
shapiro.test(prec82)
##
##  Shapiro-Wilk normality test
##
## data: prec82
## W = 0.94689, p-value = 0.2131
shapiro.test(log(prec82))
##
##  Shapiro-Wilk normality test
##
## data: log(prec82)
## W = 0.94321, p-value = 0.1755
qqnorm(prec82)
```

### Normal Q–Q Plot



```
qqnorm(log(prec82))
```

### Normal Q–Q Plot



```

preci <- na.omit(precUtm[,c('Estación', 'a1982')])
preci$a1982log <- log(preci$a1982)
preci
## Simple feature collection with 25 features and 3 fields
## geometry type:  POINT
## dimension:      XY
## bbox:            xmin: 215264.1 ymin: 1999092 xmax: 566794.7 ymax: 2197035
## epsg (SRID):   32619
## proj4string:    +proj=utm +zone=19 +datum=WGS84 +units=m +no_defs
## First 10 features:
##           Estación a1982             geom a1982log
## 1        Barahona  815.3 POINT (277900.2 2013585) 6.703556
## 2       Bayaguana 1758.6 POINT (433242.1 2073284) 7.472273
## 3        Cabrera 2136.9 POINT (405636 2171119) 7.667111
## 4       Constanza  921.2 POINT (320947.7 2090623) 6.825677
## 5 Gaspar Hernández 1844.1 POINT (363678.2 2169619) 7.519747
## 6     Hondo Valle 1709.9 POINT (215264.1 2071669) 7.444190
## 7        Jimaní  507.5 POINT (221953.7 2045651) 6.229497
## 8       La Unión 1413.1 POINT (337592.1 2184559) 7.253541
## 9        La Vega 1483.2 POINT (338847.1 2125548) 7.301957
## 10     Las Américas  787.4 POINT (429562.7 2038222) 6.668736

```

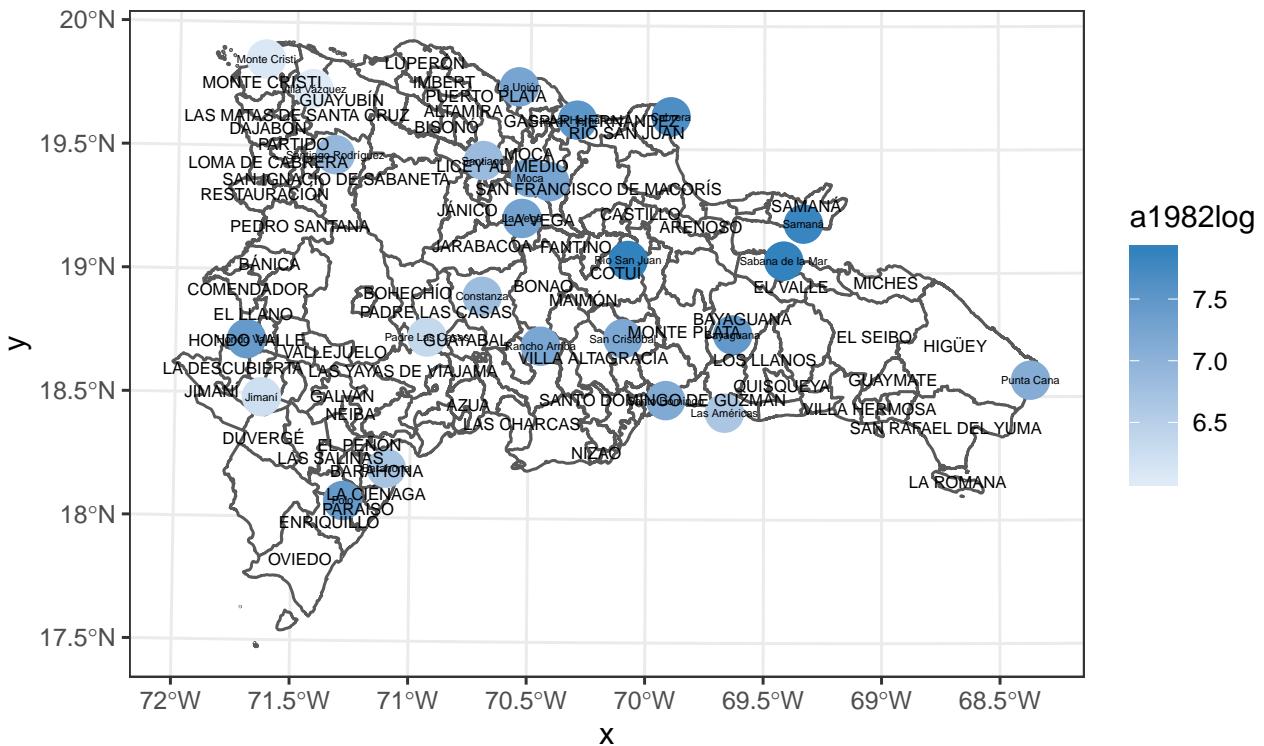
## 7.4 Despliega la localizacion de los observatorios

```

library(ggplot2)

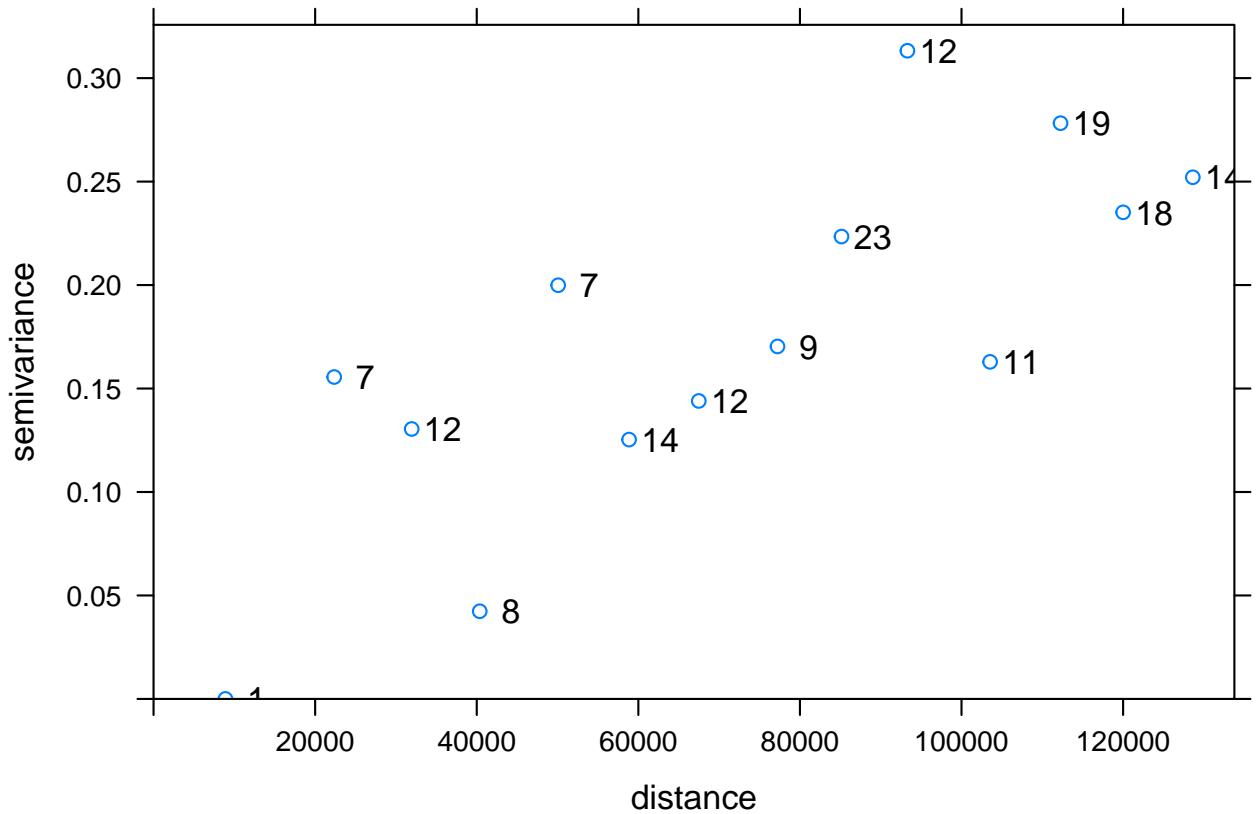
ggplot() +
  geom_sf(data = muni.sf, fill = 'white') +
  geom_sf(data = preci, aes(col = a1982log), size = 6) +
  scale_colour_gradient(low="#deebf7", high="#3182bd") +
  geom_sf_text(data = muni.sf, aes(label=TOPONIMIA), check_overlap = T, size = 2) +
  geom_sf_text(data = preci, aes(label=Estación), check_overlap = T, size = 1.5) +
  theme_bw()

```



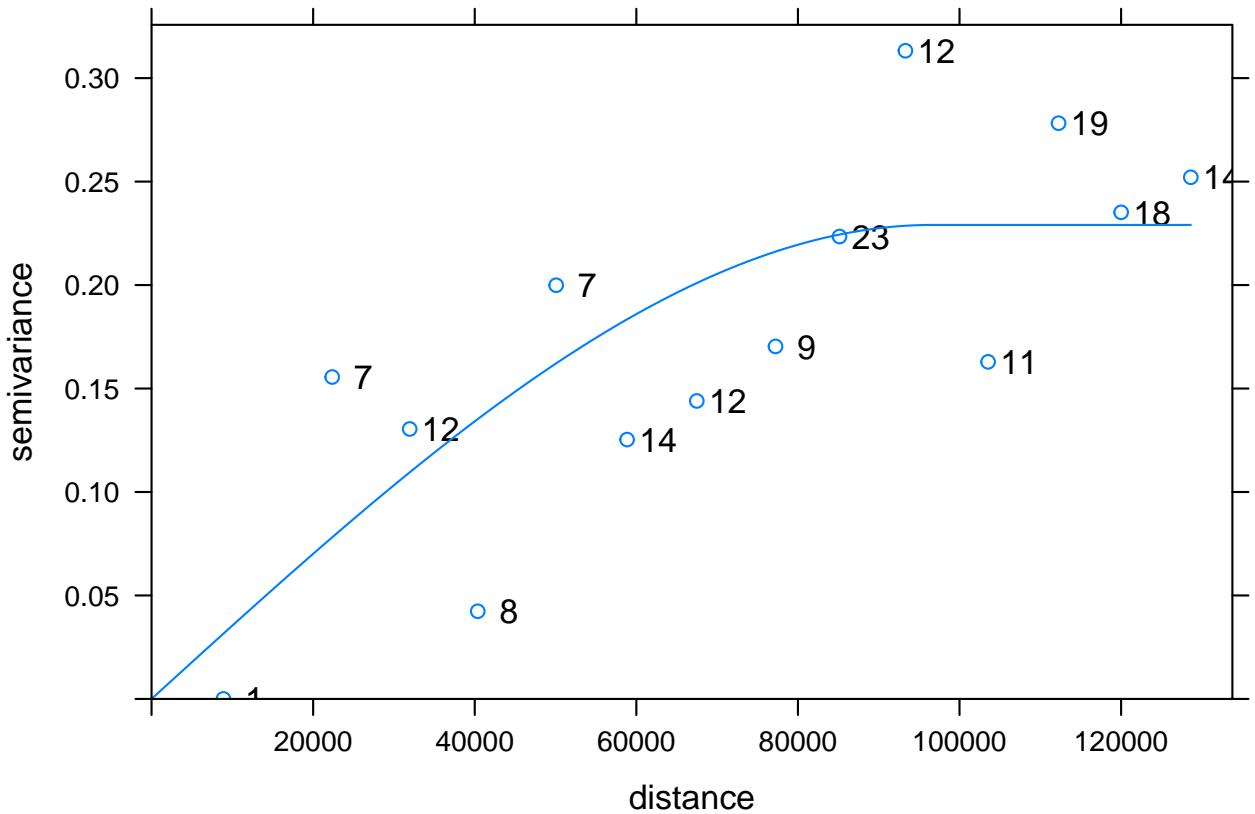
## 7.5 variograma muestral

```
v82 <- variogram(a1982log~1, preci)
v82
##   np      dist      gamma dir.hor dir.ver id
## 1  1 8896.559 1.075597e-05     0     0 var1
## 2  7 22355.182 1.555620e-01     0     0 var1
## 3 12 31950.749 1.304567e-01     0     0 var1
## 4  8 40371.385 4.236860e-02     0     0 var1
## 5  7 50078.452 1.999113e-01     0     0 var1
## 6 14 58855.498 1.253300e-01     0     0 var1
## 7 12 67482.051 1.439899e-01     0     0 var1
## 8  9 77223.824 1.703296e-01     0     0 var1
## 9 23 85146.530 2.234314e-01     0     0 var1
## 10 12 93304.363 3.132216e-01     0     0 var1
## 11 11 103535.175 1.628902e-01     0     0 var1
## 12 19 112257.676 2.782182e-01     0     0 var1
## 13 18 120004.515 2.351375e-01     0     0 var1
## 14 14 128628.216 2.520859e-01     0     0 var1
plot(v82, plot.numbers = T)
```

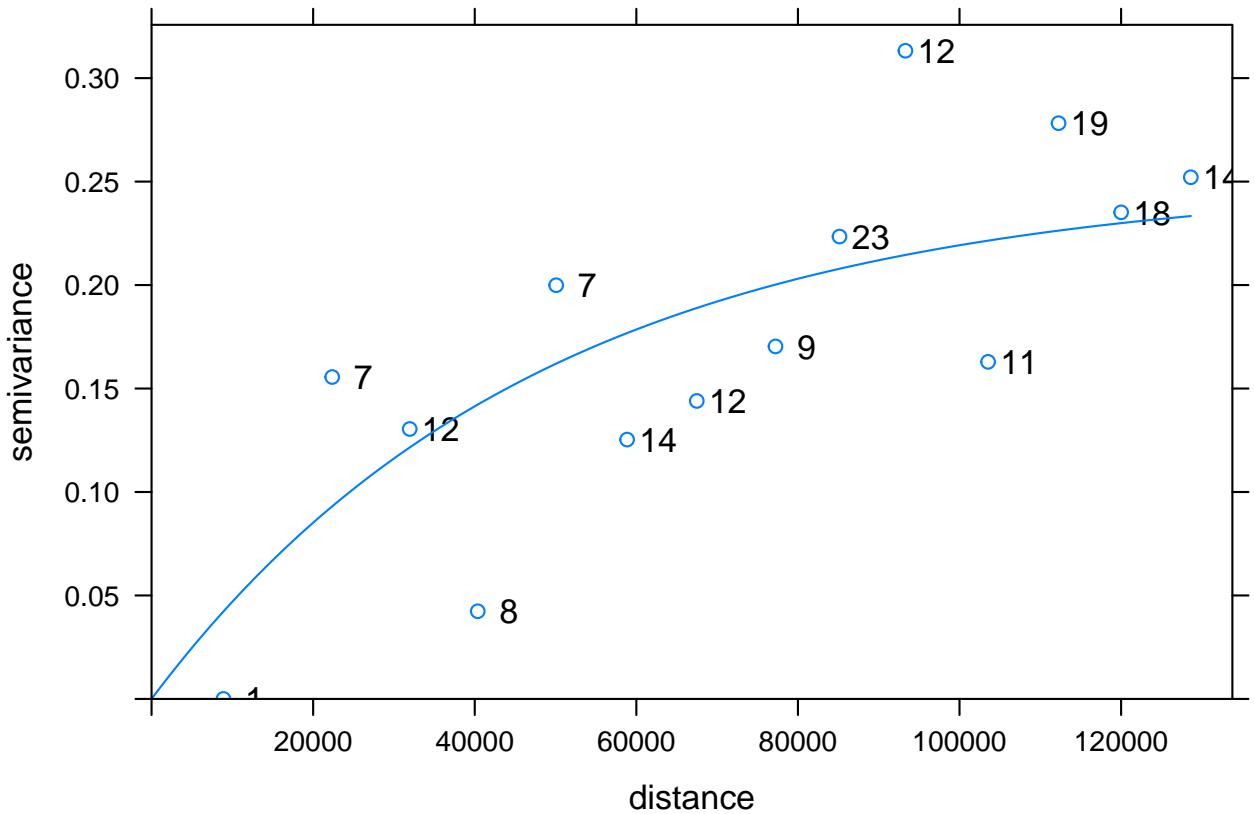


## 7.6 variograma modelo

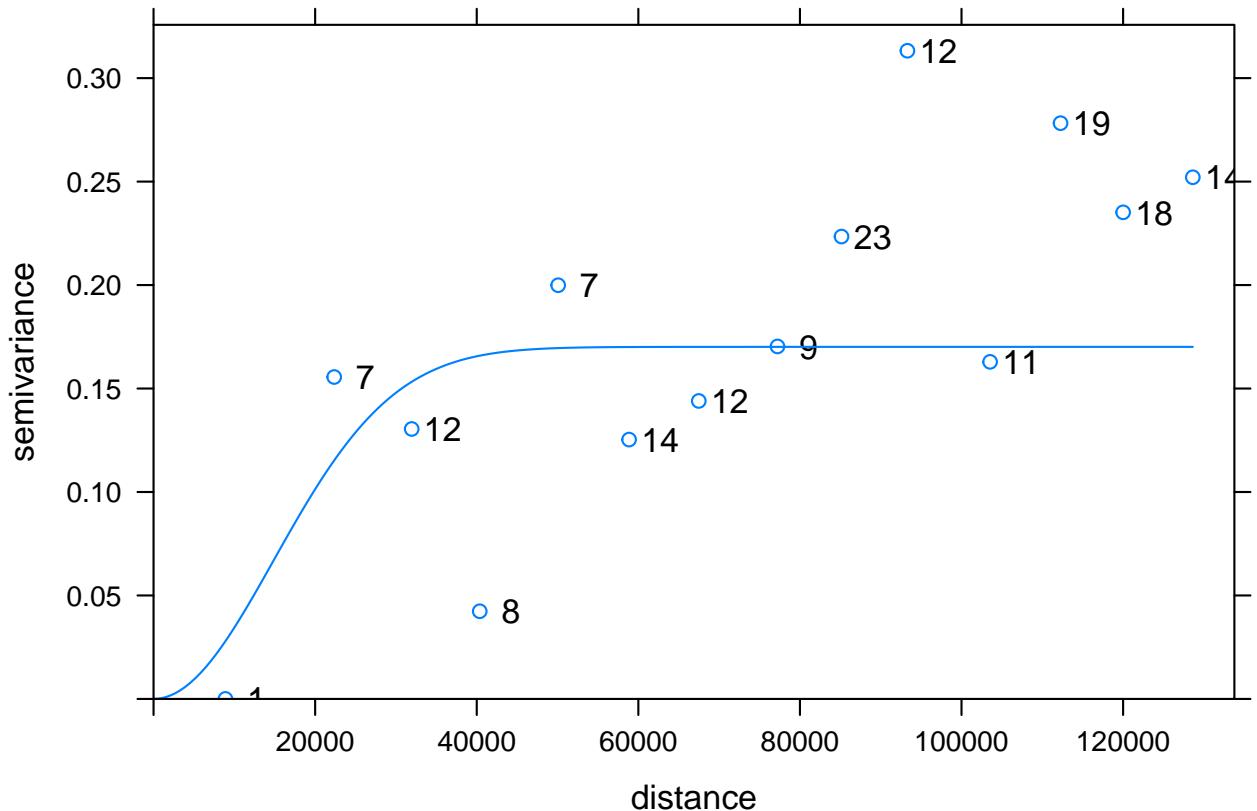
```
v82_m <- fit.variogram(v82, vgm(model = "Sph", range = 50000))
v82_m
##   model      psill      range
## 1   Sph 0.2290492 96599.97
plot(v82, v82_m, plot.numbers = T)
```



```
v82_m2 <- fit.variogram(v82, vgm(model = "Exp", range = 50000))
v82_m2
##   model      psill      range
## 1   Exp 0.2507095 48191.96
plot(v82, v82_m2, plot.numbers = T)
```



```
v82_m3 <- fit.variogram(v82, vgm(model = "Gau", range = 50000))
v82_m3
##   model      psill      range
## 1   Gau 0.1701203 21015.34
plot(v82, v82_m3, plot.numbers = T)
```



```

attr(v82_m, 'SSErr')
## [1] 1.919642e-10
attr(v82_m2, 'SSErr')                      #Elegimos el modelo exponencial
## [1] 1.706493e-10
attr(v82_m3, 'SSErr')
## [1] 1.919187e-10

```

## 7.7 kriging ordinario

```

# creacion de cuadricula de 1000
grd <- st_bbox(muni.sf) %>%
  st_as_stars(dx = 1000) %>% #10000 metros=10km de resolución espacial
  st_set_crs(crs_wgs84utm ) %>%
  st_crop(muni.sf)
grd
## stars object with 2 dimensions and 1 attribute
## attribute(s):
##   values
##   Min.    :0
##   1st Qu.:0
##   Median :0
##   Mean    :0

```

```

## 3rd Qu.:0
## Max. :0
## NA's :58018
## dimension(s):
##   from to offset delta           refsys point values
## x    1 390 182216 1000 +proj=utm +zone=19 +datum...
## y    1 272 2205216 -1000 +proj=utm +zone=19 +datum...
plot(grd)

```

**values**



```

# interpolacion
k <- krige(formula = a1982log~1, locations = preci, newdata = grd, model = v82_m2)
## [using ordinary kriging]
k
## stars object with 2 dimensions and 2 attributes
## attribute(s):
##   var1.pred      var1.var
##   Min. :6.05    Min. :0.00
##   1st Qu.:6.83   1st Qu.:0.09
##   Median :7.08   Median :0.13
##   Mean   :7.07   Mean   :0.13
##   3rd Qu.:7.28   3rd Qu.:0.16

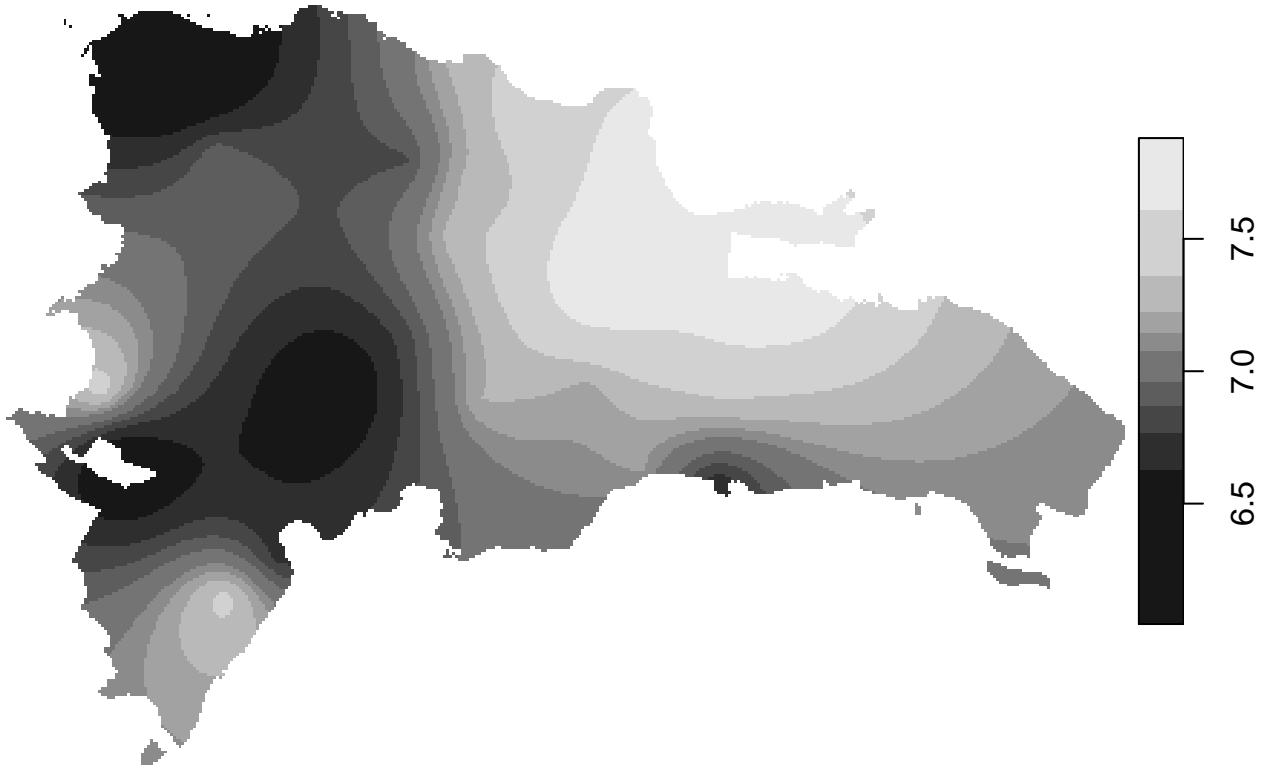
```

```

##  Max.    :7.88    Max.    :0.26
##  NA's    :58018   NA's    :58018
##  dimension(s):
##    from to offset delta                  refsys point values
##  x     1 390 182216 1000 +proj=utm +zone=19 +datum...      NA  NULL [x]
##  y     1 272 2205216 -1000 +proj=utm +zone=19 +datum...      NA  NULL [y]
plot(k)

```

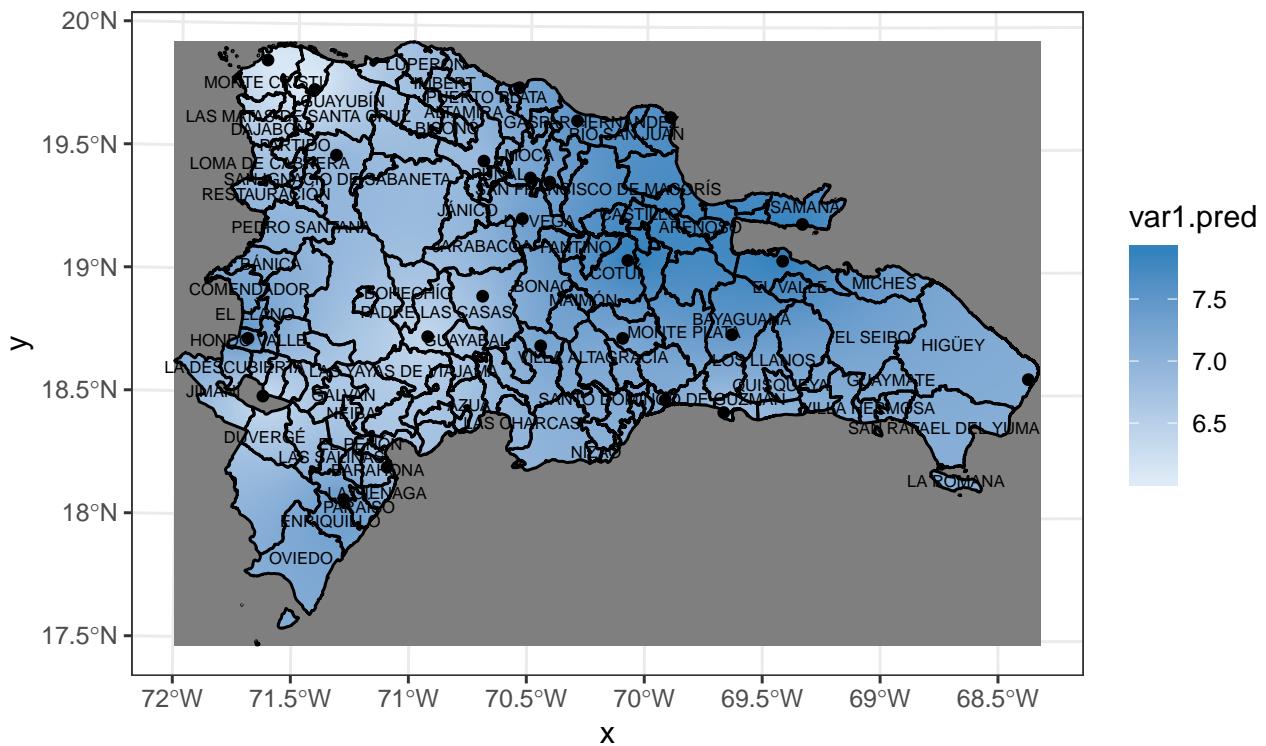
**var1.pred**



```

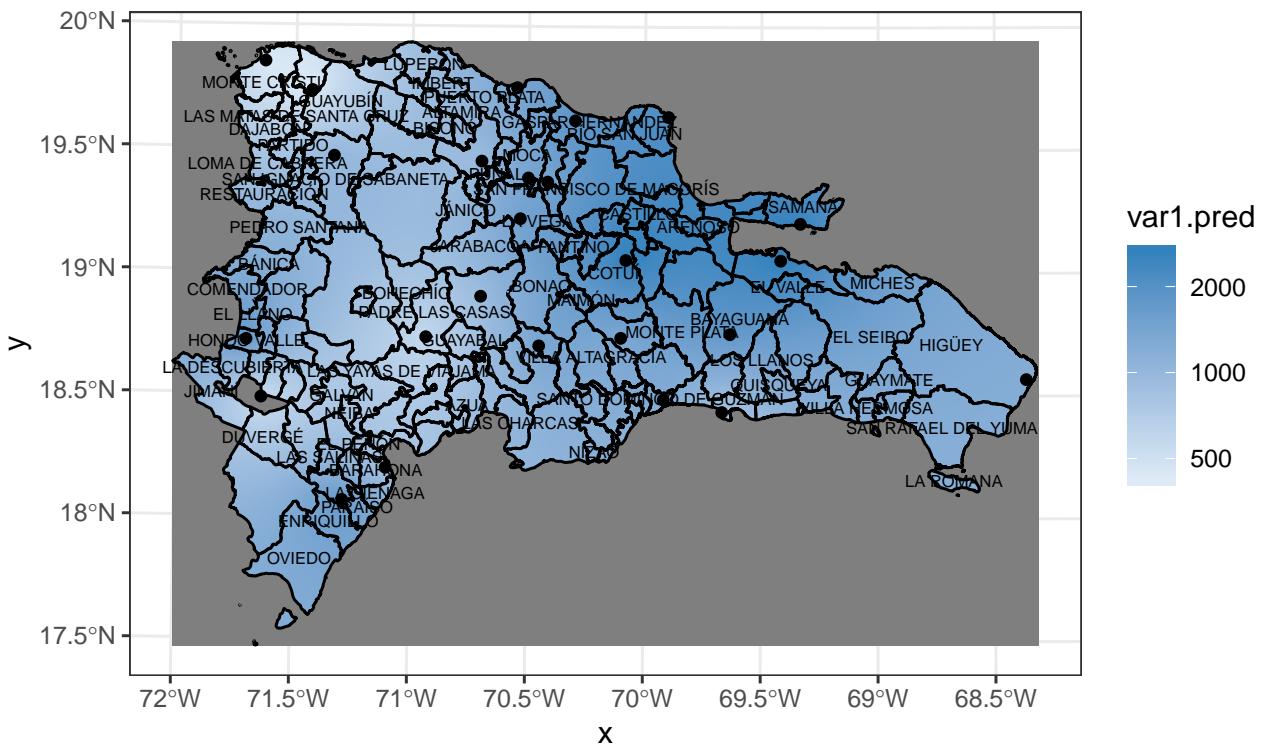
ggplot() +
  geom_stars(data = k, aes(fill = var1.pred, x = x, y = y)) +
  scale_fill_gradient(low="#deebf7", high="#3182bd") +
  geom_sf(data = st_cast(muni.sf, "MULTILINESTRING")) +
  geom_sf(data = preci) +
  geom_sf_text(data = muni.sf, aes(label=TOPONIMIA), check_overlap = T, size = 2) +
  theme_bw()

```



## 7.8 Para conseguir los valores reales de las precipitaciones

```
ggplot() +
  geom_stars(data = exp(k), aes(fill = var1.pred, x = x, y = y)) +
  scale_fill_gradient(low="#deebf7", high="#3182bd", trans = 'log10') +
  geom_sf(data = st_cast(muni_sf, "MULTILINESTRING")) +
  geom_sf(data = preci) +
  geom_sf_text(data = muni_sf, aes(label=TOPONIMIA), check_overlap = T, size = 2) +
  theme_bw()
```



## 7.9 Resultado obtenido del Análisis de precipitaciones

El análisis Kriging ordinario evidencia de una manera muy puntual aunque segregada, que los municipios localizados en la región noreste del país, recibieron una mayor cantidad de precipitaciones para el periodo correspondiente al año 1982.

## 8 FASE 3 - MODELIZACIÓN

### 8.1 Metodología

Acá se determinará la relación existente entre la variable que señala a las personas que tienen dificultad para caminar o subir escalones en el país y las de nivel educativo más alto al que asistió. Para esto se cargan las variables a relacionar de la base de datos del censo. Se calculan los valores porcentuales con respecto a la población total del país para 2010 y su valores logarítmicos. Y a partir de estos valores, se construye el modelo lineal y el modelo espacial autorregresivo.

### 8.2 Cargado de los datos

```
varsel <- censoT %>% dplyr::select(
  Toponimia = TOPONIMIA,
  PoblTotal = "Población total",
  PersDif = "Dificultad para Caminar o subir escalones: Si",
  Noasistio = "Asiste o asistió a la escuela: Nunca asistió",
  Preprimaria = "Nivel educativo más alto al que asistió: Preprimaria",
```

```

Primaria = "Nivel educativo más alto al que asistió: Primaria o básica",
Secundaria = "Nivel educativo más alto al que asistió: Secundaria o media",
Universidad = "Nivel educativo más alto al que asistió: Universitaria o superior")
varsel
## Simple feature collection with 155 features and 8 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:            xmin: 182215.8 ymin: 1933532 xmax: 571365.3 ymax: 2205216
## epsg (SRID):   32619
## proj4string:    +proj=utm +zone=19 +datum=WGS84 +units=m +no_defs
## First 10 features:
##           Toponimia PoblTotal PersDif Noasistio Preprimaria
## 1  SANTO DOMINGO DE GUZMÁN     965040   31271    45138     73199
## 2                 AZUA       91345    2616    12617     7718
## 3             LAS CHARCAS    11243     323    2329      965
## 4  LAS YAYAS DE VIAJAMA    17620     770    3442     1400
## 5        PADRE LAS CASAS   20041     804    4336     2027
## 6             PERALTA     15257     320    3528     1360
## 7          SABANA YEGUA    19020     675    2813     1079
## 8          PUEBLO VIEJO    11235     266    1629      961
## 9        TÁBARA ARRIBA    17647     437    3860     1344
## 10        GUAYABAL      5263     234     909      240
##   Primaria Secundaria Universidad           geom
## 1  290061    248102    257590 MULTIPOLYGON (((405218.1 20...
## 2  37701     19003     8601 MULTIPOLYGON (((319065.3 20...
## 3  5361      1616      204 MULTIPOLYGON (((341415.3 20...
## 4  8439      2757      541 MULTIPOLYGON (((304058.1 20...
## 5  8458      3104     1075 MULTIPOLYGON (((312890.8 20...
## 6  5736      2823     1020 MULTIPOLYGON (((317370.6 20...
## 7  9246      3676     1048 MULTIPOLYGON (((306745.8 20...
## 8  5270      2133      484 MULTIPOLYGON (((310447.9 20...
## 9  7916      2869      602 MULTIPOLYGON (((306556.7 20...
## 10 2458       924      479 MULTIPOLYGON (((322129.5 20...

```

8.3 Relativización de los datos con el campo del total de Población. Se genera las columnas de porcentaje de los datos y al mismo tiempo los logaritmos de los porcentajes encontrados.

```

varselllog <- varsel %>% mutate_each(
  funs(PCT=round(./PoblTotal,4)*100,
       PCTLLOG=log1p(round(./PoblTotal,4)*100)),
  -1, -2, -geom, -PoblTotal)
## Warning: funs() is soft deprecated as of dplyr 0.8.0
## Please use a list of either functions or lambdas:
##
##   # Simple named list:

```

```

##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(.., trim = .2), ~ median(.., na.rm = TRUE))
## This warning is displayed once per session.
varselllog
## Simple feature collection with 155 features and 20 fields
## geometry type: MULTIPOLYGON
## dimension: XY
## bbox: xmin: 182215.8 ymin: 1933532 xmax: 571365.3 ymax: 2205216
## epsg (SRID): 32619
## proj4string: +proj=utm +zone=19 +datum=WGS84 +units=m +no_defs
## First 10 features:
##           Toponimia PoblTotal PersDif Noasistio Preprimaria
## 1 SANTO DOMINGO DE GUZMÁN    965040    31271    45138    73199
## 2                      AZUA    91345     2616    12617    7718
## 3                      LAS CHARCAS    11243     323    2329    965
## 4          LAS YAYAS DE VIAJAMA    17620     770    3442    1400
## 5          PADRE LAS CASAS    20041     804    4336    2027
## 6                      PERALTA    15257     320    3528    1360
## 7                      SABANA YEGUA    19020     675    2813    1079
## 8          PUEBLO VIEJO    11235     266    1629     961
## 9          TÁBARA ARRIBA    17647     437    3860    1344
## 10                     GUAYABAL    5263     234     909     240
##           Primaria Secundaria Universidad
## 1      290061     248102    257590 MULTIPOLYGON (((405218.1 20...
## 2      37701      19003     8601 MULTIPOLYGON (((319065.3 20...
## 3      5361       1616     204 MULTIPOLYGON (((341415.3 20...
## 4      8439       2757     541 MULTIPOLYGON (((304058.1 20...
## 5      8458       3104     1075 MULTIPOLYGON (((312890.8 20...
## 6      5736       2823     1020 MULTIPOLYGON (((317370.6 20...
## 7      9246       3676     1048 MULTIPOLYGON (((306745.8 20...
## 8      5270       2133     484 MULTIPOLYGON (((310447.9 20...
## 9      7916       2869     602 MULTIPOLYGON (((306556.7 20...
## 10     2458       924      479 MULTIPOLYGON (((322129.5 20...
##           PersDif_PCT Noasistio_PCT Preprimaria_PCT Primaria_PCT Secundaria_PCT
## 1        3.24        4.68       7.59      30.06      25.71
## 2        2.86       13.81       8.45      41.27      20.80
## 3        2.87       20.72       8.58      47.68      14.37
## 4        4.37       19.53       7.95      47.89      15.65
## 5        4.01       21.64      10.11      42.20      15.49
## 6        2.10       23.12       8.91      37.60      18.50
## 7        3.55       14.79       5.67      48.61      19.33
## 8        2.37       14.50       8.55      46.91      18.99

```

```

## 9      2.48      21.87      7.62      44.86      16.26
## 10     4.45      17.27      4.56      46.70      17.56
## Universidad_PCT PersDif_PCTLOG Noasisstio_PCTLOG Preprimaria_PCTLOG
## 1      26.69      1.444563      1.736951      2.150599
## 2      9.42      1.350667      2.695303      2.246015
## 3      1.81      1.353255      3.078233      2.259678
## 4      3.07      1.680828      3.021887      2.191654
## 5      5.36      1.611436      3.119718      2.407846
## 6      6.69      1.131402      3.183041      2.293544
## 7      5.51      1.515127      2.759377      1.897620
## 8      4.31      1.214913      2.740840      2.256541
## 9      3.41      1.247032      3.129826      2.154085
## 10     9.10      1.695616      2.905260      1.715598
## Primaria_PCTLOG Secundaria_PCTLOG Universidad_PCTLOG
## 1      3.435921      3.285038      3.321071
## 2      3.744078      3.081910      2.343727
## 3      3.885268      2.732418      1.033184
## 4      3.889573      2.812410      1.403643
## 5      3.765840      2.802754      1.850028
## 6      3.653252      2.970414      2.039921
## 7      3.904192      3.012098      1.873339
## 8      3.869324      2.995232      1.669592
## 9      3.825593      2.848392      1.483875
## 10     3.864931      2.921009      2.312535

```

Ya habíamos evaluado la autocorrelación, la normalidad y la homocedasticidad de nuestra variable dependiente “Personas con dificultad para caminar o subir escalones” en la primera fase de este análisis.

#### 8.4 Construcción de un modelo lineal

```

modlin <- varsellog %>% select(contains('_PCTLOG')) %>%
  st_drop_geometry() %>% lm(PersDif_PCTLOG ~ ., .)
modlin %>% summary
##
## Call:
## lm(formula = PersDif_PCTLOG ~ ., data = .)
##
## Residuals:
##       Min     1Q   Median     3Q    Max 
## -0.56333 -0.09969 -0.00397  0.10800  0.41231
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -5.75228   1.32333 -4.347 2.55e-05 ***  
## Noasisstio_PCTLOG  0.24309   0.09668  2.514   0.013 *  
## 
```

```

## Preprimaria_PCTLOG -0.02990  0.09523 -0.314   0.754
## Primaria_PCTLOG     1.51875  0.19163  7.926 4.91e-13 ***
## Secundaria_PCTLOG    0.08914  0.13416  0.664   0.507
## Universidad_PCTLOG   0.34383  0.06500  5.289 4.30e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ',' 1
##
## Residual standard error: 0.1698 on 149 degrees of freedom
## Multiple R-squared:  0.3562, Adjusted R-squared:  0.3346
## F-statistic: 16.49 on 5 and 149 DF,  p-value: 6.149e-13
modlin %>% bptest
##
## studentized Breusch-Pagan test
##
## data: .
## BP = 8.7067, df = 5, p-value = 0.1214

```

## 8.5 Construcción del modelo espacial autorregresivo

```

sar <- varselllog %>% select(contains('_PCTLOG')) %>%
  st_drop_geometry() %>%
  spautolm(formula = PersDif_PCTLOG ~ ., data = ., listw = censo.w.w)
summary(sar)
##
## Call: spautolm(formula = PersDif_PCTLOG ~ ., data = ., listw = censo.w.w)
##
## Residuals:
##       Min        1Q      Median        3Q       Max
## -0.527208 -0.101207 -0.010725  0.103898  0.417982
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.936618  1.410323 -3.5003 0.0004647
## Noasistio_PCTLOG  0.133800  0.106651  1.2546 0.2096355
## Preprimaria_PCTLOG -0.014874  0.090198 -0.1649 0.8690190
## Primaria_PCTLOG    1.416742  0.201758  7.0220 2.187e-12
## Secundaria_PCTLOG   0.095323  0.145183  0.6566 0.5114537
## Universidad_PCTLOG  0.255329  0.067148  3.8025 0.0001433
##
## Lambda: 0.47831 LR test value: 19.374 p-value: 1.0745e-05
## Numerical Hessian standard error of lambda: 0.097706
##
## Log likelihood: 67.6154
## ML residual variance (sigma squared): 0.023227, (sigma: 0.1524)
## Number of observations: 155

```

```

## Number of parameters estimated: 8
## AIC: -119.23

sar2 <- varsellog %>% select(contains('PCTLOG')) %>%
  st_drop_geometry() %>%
  spautolm(formula = PersDif_PCTLOG ~ Noasistio_PCTLOG + Primaria_PCTLOG + Universidad_PCTLOG, d
summary(sar2)
##
## Call:
## spautolm(formula = PersDif_PCTLOG ~ Noasistio_PCTLOG + Primaria_PCTLOG +
##           Universidad_PCTLOG, data = ., listw = censo.w.w)
##
## Residuals:
##       Min      1Q   Median      3Q     Max
## -0.531224 -0.101392 -0.012913  0.104317  0.417114
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)             -4.479750  0.847949 -5.2830 1.271e-07
## Noasistio_PCTLOG        0.087980  0.069897  1.2587  0.2081
## Primaria_PCTLOG         1.395665  0.172588  8.0867 6.661e-16
## Universidad_PCTLOG      0.256877  0.063985  4.0147 5.953e-05
##
## Lambda: 0.47855 LR test value: 19.534 p-value: 9.884e-06
## Numerical Hessian standard error of lambda: 0.097297
##
## Log likelihood: 67.35007
## ML residual variance (sigma squared): 0.023305, (sigma: 0.15266)
## Number of observations: 155
## Number of parameters estimated: 6
## AIC: -122.7

```

## 8.6 Resultados obtenidos

Las variables significativas encontradas al realizar el modelo de regresión lineal son: El nunca asistió a la escuela (Noasistió), y los niveles educativos de Primaria y Universitaria. Dado que obtuvimos un valor de p mayor a 0.05, concluimos que no se cumple la propiedad de homocedasticidad. Presenta heterocedasticidad.

El coeficiente de regresión de la variable dependiente en el modelo autorregresivo es negativo, lo que indica una relación inversa.

Para estos datos obtenidos podríamos deducir que hay relación directa entre el grupo de personas que no pueden caminar y las que terminaron sus estudios de primaria y universitaria: Para cada 1.4% del logaritmo de porcentaje de personas cuyo nivel educativo más alto es el de Primaria, solo se aumenta en 0.17% el número de personas con dificultad para caminar o subir escalones. Y en el caso del 0.26% del logaritmo de porcentaje de personas cuyo nivel educativo más alto es el universitario, solo aumenta en 0.06% el número de personas con dificultad para caminar o subir

escalones. Y que la variable de Nunca asistió a la escuela no es realmente significativa.