589Project_b

Shengyao Li

2024-04-29

```
install.packages("data.table")
```

```
##
## The downloaded binary packages are in
##  /var/folders/3b/l64jrnxs3zb4gqrm_qnntv780000gn/T//Rtmpi9GzQE/downloaded_p
ackages
```

```
library(data.table)
dataset <- fread("589_RawData.csv")
```

```
library(dplyr)
```

```
locality_selected_data <- dataset %>% select(gbifID, decimalLatitude, localit
y, decimalLongitude, eventDate, verbatimScientificName)
```

```
write.csv(locality_selected_data, file = "locality_selected_data.csv", row.na
mes = FALSE)
```

```
library(dplyr)
```

```
locality_counts <- locality_selected_data %>%
  group_by(locality) %>%
  summarise(count = n())
```

```
print(locality_counts)
```

```
## # A tibble: 239 × 2
##    locality                       count
##    <chr>                          <int>
##  1 ""                                59
##  2 "1001 steps to Beecher Street"   154
##  3 "131st Street to 1001 steps"     164
##  4 "Aguilar Point"                   11
##  5 "Albert Head Lagoon"              96
##  6 "Ambleside"                      171
##  7 "Ardmore/Coles Bay"               66
##  8 "Barnet East"                     32
##  9 "Barnet West"                     91
## 10 "Batchelor Pt. - Gleneagles"      71
## # ℹ 229 more rows
```

```
write.csv(locality_counts, file = "locality_counts.csv", row.names = FALSE)
```

```r
count_by_coordinates <- locality_selected_data %>%
  group_by(decimalLatitude, decimalLongitude) %>%
  summarise(count = n(), .groups = 'drop')

print(count_by_coordinates)
```

```
## # A tibble: 203 × 3
##    decimalLatitude decimalLongitude count
##              <dbl>            <dbl> <int>
##  1            48.3            -124.   243
##  2            48.4            -124.    27
##  3            48.4            -123.    96
##  4            48.4            -123.   227
##  5            48.4            -123.   141
##  6            48.4            -123.   147
##  7            48.4            -123.   119
##  8            48.4            -123.    57
##  9            48.4            -123.    42
## 10            48.4            -123.    86
## # ℹ 193 more rows
```

```r
write.csv(count_by_coordinates, file = "count_by_coordinates.csv", row.names
= FALSE)
```

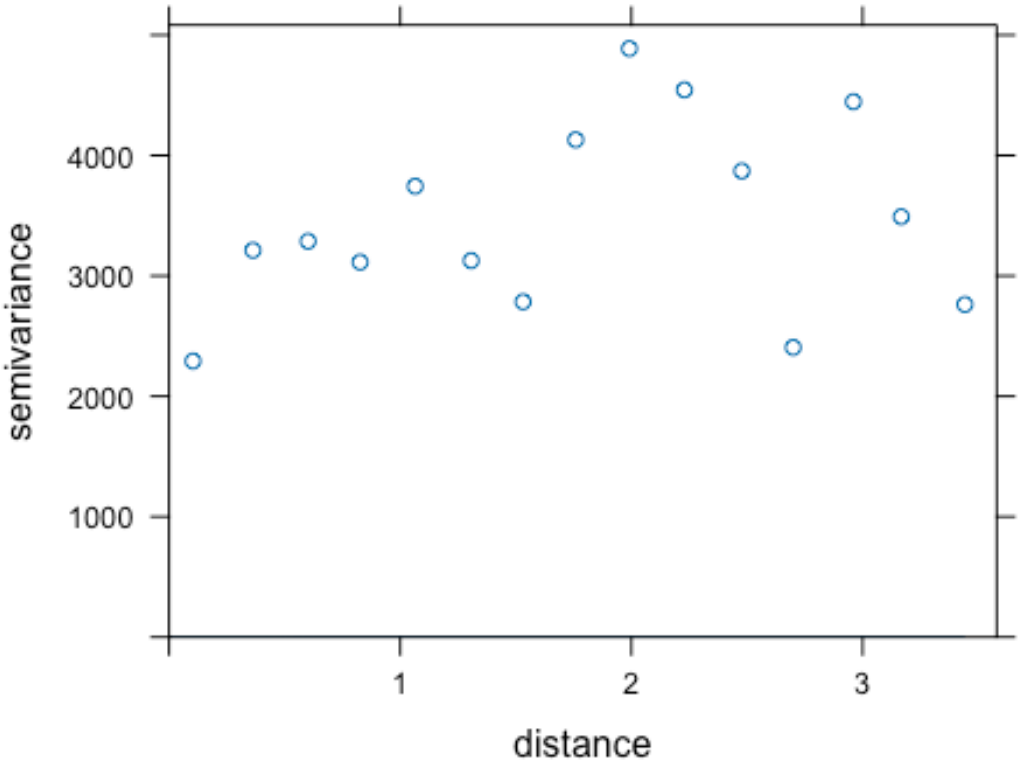
Variogram

```r
install.packages(c("gstat", "sp"))
```

```
##
## The downloaded binary packages are in
##  /var/folders/3b/l64jrnxs3zb4gqrm_qnntv780000gn/T//Rtmpi9GzQE/downloaded_p
ackages
```

```r
library(gstat)
library(sp)


count_by_coordinates <- na.omit(count_by_coordinates)

coordinates <- count_by_coordinates[, c("decimalLatitude", "decimalLongitude"
)]
counts <- count_by_coordinates$count

coordinates_df <- as.data.frame(coordinates)

coordinates_sp <- SpatialPoints(coordinates_df)

spdf <- SpatialPointsDataFrame(coordinates_sp, data=data.frame(count=counts))
```

```r
v <- variogram(count~1, spdf)

fit <- fit.variogram(v, model=vgm(psill=1, "Sph", range=1, nugget=0.5))

## Warning in fit.variogram(v, model = vgm(psill = 1, "Sph", range = 1, nugge
t =
## 0.5)): No convergence after 200 iterations: try different initial values?

plot(v, fit)
```




```r
print(fit)

##   model     psill     range
## 1   Nug 0.5285531 0.0000000
## 2   Sph 1.0193704 0.9887494
```

Kriging

```r
grid_size <- 0.01

grd <- expand.grid(
  decimalLatitude = seq(from = min(count_by_coordinates$decimalLatitude), to
= max(count_by_coordinates$decimalLatitude), by = grid_size),
```

```
  decimalLongitude = seq(from = min(count_by_coordinates$decimalLongitude), t
o = max(count_by_coordinates$decimalLongitude), by = grid_size)
)

if (!nrow(grd) > 0) {
  stop("The grid has no rows. Check the sequence generation for latitude and
longitude.")
}

coordinates(grd) <- ~decimalLatitude+decimalLongitude
gridded(grd) <- TRUE

grd_sp <- SpatialPixelsDataFrame(grd, data=data.frame(count=rep(NA, length(gr
d))))

krige_out <- krige(formula = count~1, locations = spdf, newdata = grd_sp, mod
el = fit)

## [using ordinary kriging]

spplot(krige_out, "var1.pred",
       scales = list(draw = TRUE))
```