Making maps and analysing spatial data: An introduction to using R for archaeologists／考古学者のためのRによる地図作成と空間データ分析

Liying Wang

Atsushi Noguchi (translator)

6/14/2020

# Introduction／はじめに

In this workshop, you will learn how to use R to manipulate and visualize spatial data of interest to archaeologists. This rmd file is a demonstration of code and you will work on it step by step to get familiar with spatial data and basic analysis. There are three main topics covered in this workshop:

-Part 1: making maps, including regional map and site map -Part 2: spatial data manipulation and visualization -Part 3: spatial data analysis, including density analysis of sites, and hypothesis testing for their distributions

このワークショップでは、Rを用いて、考古学者にとっての関心のある空間データを操作し可視化する方法を学びます。このRマークダウン(.rmd)ファイルは、Rのコードを実行し、参加者が手順を追って空間データと基礎的な分析を理解するためのものです。以下の3つの主要なトピックを含みます：

-Part 1: 地図の作成- 地域地図と遺跡地図  
-Part 2: 空間データの操作と可視化  
-Part 3: 空間データ分析- 遺跡分布密度分析と遺跡分布に関する仮説検証

Before getting started, make sure you have all packages we will use! If not, run the code below:

開始する前に、必要なパッケージがすべて揃っているかどうか確認しましょう。もし揃っていなければ、以下のコードを実行します。

## Set up 1: install packages／セットアップ1：パッケージのインストール

Copy and paste the code to your console and run it: install.packages(c(“rnaturalearth”, “rnaturalearthdata”, “ggplot2”, “tidyverse”, “sf”, “sp”,“shadowtext”, “ggmap”, “ggspatial”, “raster”, “spatstat”, “maptools”))

上記のコードをコピー、コンソールにペーストして実行しましょう。

If you see a message “Do you want to install from sources the package which needs compilation?” Type “No” on your console, press Enter, and it will continue to download.

もし“Do you want to install from sources the package which needs compilation?”というメッセージが表示されたら、“No”と入力し、エンターキーを押してください。必要なパッケージのダウンロードが継続します。

Copy and paste the code to your console and run it: devtools::install\_github(‘3wen/legendMap’)

ダウンロードとインストールが完了したら、上記のコードをコンソールにコピー・ペーストして実行しましょう。 （パッケージのインストールが完了している場合は不要です）

## Set up 2: create data folder to store spatial data for this workshop／セットアップ2：ワークショップ用の空間データを保存するフォルダの作成

# create data folder to store the raster data ／ラスター・データを保存するためのフォルダ作成  
dir.create("data")

## Warning in dir.create("data"): 'data' already exists

# download the raster zip file into our data folder／ラスターzipファイルのダウンロード  
download.file("https://github.com/LiYingWang/Japan\_GIS\_workshop\_202006\_LW/raw/master/workshop\_data.zip", "data/raster-shapefile.zip")  
  
# unzip to the data folder／データフォルダに解凍  
unzip(zipfile = "data/raster-shapefile.zip", exdir = "data")  
  
# delete zip file／不要な.zipファイルの削除  
unlink("data/raster-shapefile.zip")

# Making maps／地図の作成

Load world data and take a look at the data form, especially the “geometry” column where it stores the coordinates we need for making maps.

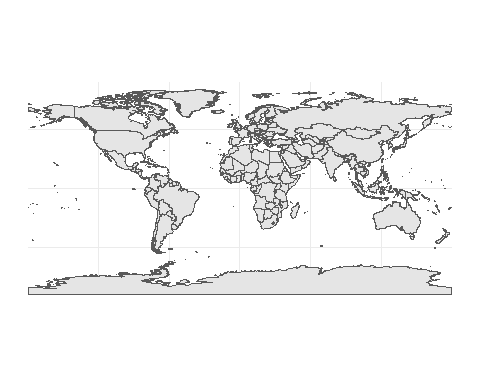
世界地図のデータを読み込み、データの形式、とくに地図作成に必要な座標が収納されている“geometry”カラムを確認します。

## Exercise 1: load data and plot the world map (5 mins)／実習1：データの読み込みと世界地図の描画

library(rnaturalearth) # provides world map／世界地図を描画するrnaturalearthパッケージをアクティベート  
library(rnaturalearthdata) #arnaturalearthdataパッケージをアクティベート  
  
world <- ne\_countries(scale = "medium", returnclass = "sf") # pulls country data／国データを取得  
class(world) # what class it is?／ここでのオブジェクト・クラスは何ですか?

## [1] "sf" "data.frame"

# type View(world) in your console to take a look at the data frame  
# コンソールに `View(world)` と入力し、データフレームを確認しましょう  
  
library(ggplot2) #ggplot2のアクティベート  
  
# plot basic world map／基本的な世界地図の描画   
ggplot(data = world) +  
 geom\_sf() + # adds a geometry stored in world／オブジェクトworldに収納されたジオメトリを描画に追加  
 theme\_minimal()



Now, we want to plot Japan with some countries around it as our regional map. We may want to indicate the countries by adding name label on it. To do this, we need to get the center of the country for adding country labels, and then specify which countries we want to show their names on the map.

ここでは、日本を周辺のいくつかの国々とともに地域地図として描画し、国名を表示させたいと思います。そのために、中央に表示する国を指定し、表示範囲に含む国々を特定します。

## Exercise 2: make a regional map (7 mins)／実習2：地域地図の作成

library(tidyverse) #tidyverseのアクティベート

## -- Attaching packages ----------------------------------------------------------------------- tidyverse 1.3.0 --

## v tibble 3.0.1 v dplyr 0.8.5  
## v tidyr 1.0.2 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0  
## v purrr 0.3.4

## -- Conflicts -------------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

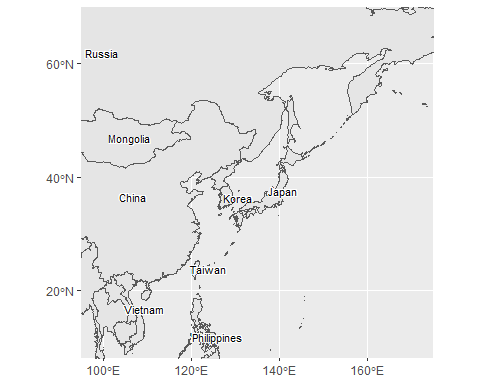
library(sf) #sf(Simple Features)のアクティベート ※空間データを扱うため

## Linking to GEOS 3.8.0, GDAL 3.0.4, PROJ 6.3.1

country\_centre\_coords <-  
 as\_tibble(st\_coordinates(st\_centroid(world$geometry))) # for the text labels／国名表示用

## Warning in st\_centroid.sfc(world$geometry): st\_centroid does not give correct  
## centroids for longitude/latitude data

world\_points <-  
 world %>%  
 bind\_cols(country\_centre\_coords) %>%  
 filter(name %in% c("Japan", "China", "Korea", "Taiwan", "Russia",  
 "Philippines", "Vietnam", "Mongolia")) # 表示に含める国々をworldオブジェクトから抽出  
  
library(shadowtext) #shadowtextのアクティベート ※ggplot2にshadowtextレイヤーを追加するため  
  
# plot map／マップの描画  
JP\_NE\_Asia <-  
 ggplot(data = world) +  
 geom\_sf() +  
 geom\_shadowtext(data= world\_points, # add texts／国名の追加   
 aes(x = X, y = Y,  
 label = name),  
 color='black',  
 bg.colour='white',  
 size = 3,  
 position = position\_nudge(y = 0, x = 3)) +  
 coord\_sf(xlim = c(95, 175), # zoom in the area of interest／地図表示範囲の指定(x=経度,y=緯度)   
 ylim = c(8, 70),   
 expand = FALSE) + # match the limits we provide／指定範囲で切り抜き   
 scale\_x\_continuous(breaks = seq(100, 160, by = 20)) + #経緯線の間隔指定  
 scale\_y\_continuous(breaks = seq(20, 65, by = 20)) +  
 theme(axis.title.x = element\_blank(),  
 axis.title.y = element\_blank())  
  
JP\_NE\_Asia



We also want to make a map with archaeological sites we are interested in. Since we would add points for sites, we have to specify the coordinates of site location first.

次に考古学遺跡の位置も追加したいと思います。そのためにはまず、遺跡の位置の座標を指定しなければなりません。

## Exercise 3: make site map (7 mins)

# add site location／遺跡位置の追加   
site\_location <-  
 data.frame(location = c("Daisen Kofun", "Todai-ji temple"),  
 lon = c(135.487953, 135.839891),  
 lat = c(34.564503, 34.688862)) #Daisen Kofun(大山古墳)とTodai-ji temple(東大寺)の位置情報をsite\_locationに収納  
  
library(ggmap) #ggmapのアクティベート ※Rで地図表示を行なうため

## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.

## Please cite ggmap if you use it! See citation("ggmap") for details.

library(ggspatial) #ggspatialのアクティベート ※ggplot2で地図・空間データを描画するため  
library(legendMap) #legendmapのアクティベート ※ggplot2で方位・スケールを描画するため

## Loading required package: maps

##   
## Attaching package: 'maps'

## The following object is masked from 'package:purrr':  
##   
## map

## Loading required package: maptools

## Loading required package: sp

## Checking rgeos availability: TRUE

## Loading required package: grid

local\_map <- ggmap(get\_stamenmap(rbind(as.numeric(c(135.3, 34.3,  
 136.0, 35))), zoom = 10)) # define the range／地域地図の描画範囲(緯度経度で指定)とズームレベルを指定

## Source : http://tile.stamen.com/terrain/10/896/405.png

## Source : http://tile.stamen.com/terrain/10/897/405.png

## Source : http://tile.stamen.com/terrain/10/898/405.png

## Source : http://tile.stamen.com/terrain/10/896/406.png

## Source : http://tile.stamen.com/terrain/10/897/406.png

## Source : http://tile.stamen.com/terrain/10/898/406.png

## Source : http://tile.stamen.com/terrain/10/896/407.png

## Source : http://tile.stamen.com/terrain/10/897/407.png

## Source : http://tile.stamen.com/terrain/10/898/407.png

## Source : http://tile.stamen.com/terrain/10/896/408.png

## Source : http://tile.stamen.com/terrain/10/897/408.png

## Source : http://tile.stamen.com/terrain/10/898/408.png

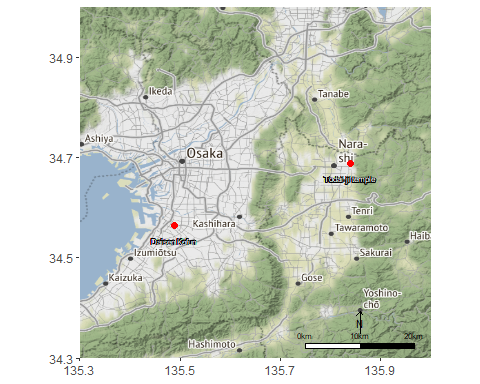
site\_Japan <-   
 local\_map +  
 geom\_point(data = site\_location, # add a layer of sites／遺跡のレイヤーを追加   
 aes(x = lon,  
 y = lat),  
 size = 2,  
 color = "red") +  
 geom\_shadowtext(data = site\_location, # add texts   
 aes(x = lon,  
 y = lat,  
 label = location),  
 size = 2,  
 position = position\_nudge(y = - 0.03),  
 check.overlap = TRUE) +  
 coord\_sf(xlim = c(135.3, 136), # define the range／描画範囲を定義   
 ylim = c(34.3, 35),  
 expand = FALSE) +  
 scale\_x\_continuous(breaks = seq(135.3, 136, by = 0.2)) +   
 scale\_y\_continuous(breaks = seq(24.3, 35, by = 0.2)) + #経緯線の間隔を指定   
 legendMap::scale\_bar(   
 lon = 135.75,  
 lat = 34.32,  
 legend\_size = 2,  
 # distance of one section of scale bar, in km  
 distance\_lon = 10,  
 # height of the scale bar, in km  
 distance\_lat = 1,  
 # distance between scale bar and units, in km  
 distance\_legend = 3,  
 # units of scale bar  
 dist\_unit = "km",  
 # add the north arrow  
 orientation = TRUE,  
 # length of N arrow, in km  
 arrow\_length = 5,  
 # distance between scale bar & base of N arrow, in km  
 arrow\_distance = 3,  
 # size of letter 'N' on N arrow, in km  
 arrow\_north\_size = 3) +  
 theme(axis.title.x = element\_blank(),  
 axis.title.y = element\_blank())

## Coordinate system already present. Adding new coordinate system, which will replace the existing one.

## Scale for 'x' is already present. Adding another scale for 'x', which will  
## replace the existing scale.

## Scale for 'y' is already present. Adding another scale for 'y', which will  
## replace the existing scale.

site\_Japan



（大阪と奈良を中心とした範囲の地図が描画され、大山古墳と東大寺の位置がプロットされます。もし2つの遺跡のドットとラベルが小さくて見えない場合は、165行目と171行目のsize=の数値を大きくしてください。

We can save the map using ggsave function below the plot we would like to save.

ggsave関数を利用して地図を保存することができます。

# save the map to your folder／フォルダに作成した地図を保存   
ggsave(here::here("Japan-site-map.jpg"),  
 width = 60,  
 height = 60,  
 dpi = 300,  
 units = "mm")

# Spatial data manipulation and visualization／空間データの操作と可視化

The raster data is DEM data downloaded from <https://www.gsi.go.jp/kankyochiri/gm_japan_e.html>. We want to crop the area that matches the site map from this DEM data. Here, we use coordinates to create a data frame, convert it to a spatial object, and then use it to crop the raster.

地図に使用したラスターデータは、国土地理院が提供するDEMデータです。遺跡地図で表示したい範囲のDEMデータを切り出したいと思います。位置座標でデータフレームを作成し、それを空間データオブジェクトに変換し、ラスターデータの切り抜きに使用します。

## Exercise 4: crop an area from raster data (7 mins)／実習4：ラスターデータの切り抜き

library(raster) #rasterのアクティベート ※ラスターデータ表示のため

##   
## Attaching package: 'raster'

## The following object is masked from 'package:dplyr':  
##   
## select

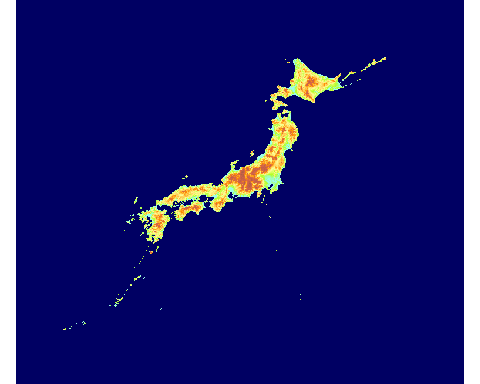
## The following object is masked from 'package:tidyr':  
##   
## extract

# read in data from data folder／フォルダからデータの読み込み  
DEM\_Japan <- raster("data/workshop\_data/jpn/el.tif") #ダウンロード済みデータをDEM\_Japanに収納 ※日本全域   
  
# assign coordinate reference system／座標系の指定  
crs(DEM\_Japan) <- "+proj=lcc +lat\_1=41.03333333333333 +lat\_2=40.66666666666666 +lat\_0=40.16666666666666 +lon\_0=-74 +x\_0=300000 +y\_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=us-ft +no\_defs"

## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO"): Discarded datum Unknown based on GRS80 ellipsoid in CRS definition,  
## but +towgs84= values preserved

plot(DEM\_Japan) # take a look at raster data／ラスターデータの描画

## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO"): Discarded datum Unknown based on GRS80 ellipsoid in CRS definition,  
## but +towgs84= values preserved  
  
## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO"): Discarded datum Unknown based on GRS80 ellipsoid in CRS definition,  
## but +towgs84= values preserved

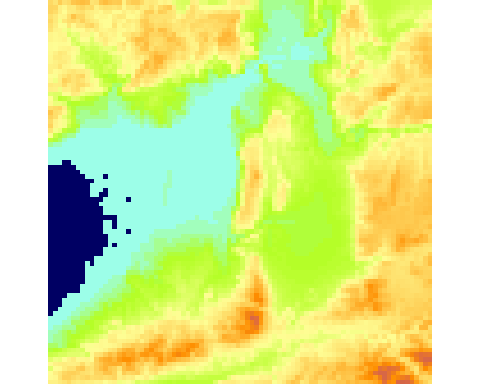


# define the area that we want to crop from the DEM／切り出したい範囲の定義   
x\_coord <- c(135.3, 135.3, 136, 136, 135.3)  
y\_coord <- c(34.3, 35, 35, 34.3, 34.3)  
xym <- cbind(x\_coord, y\_coord)  
  
library(sp) #spのアクティベート ※空間データを取り扱うため   
p = Polygon(xym) # convert the matrix to polygon／行列をポリゴンに変換   
ps = Polygons(list(p),1) # make lists／リストの作成   
sps = SpatialPolygons(list(ps)) # convert to Spatial Polygons／空間ポリゴンデータに変換   
crs(sps) <- crs(DEM\_Japan) # define coordinate reference system／参照座標系の定義   
crop\_DEM <- crop(DEM\_Japan, sps) # crop a raster to the extent of specified spatial object／設定した空間データオブジェクトによりラスターデータを切り出し

## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO"): Discarded datum Unknown based on GRS80 ellipsoid in CRS definition,  
## but +towgs84= values preserved

plot(crop\_DEM) #切り出したラスターデータの描画

## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO"): Discarded datum Unknown based on GRS80 ellipsoid in CRS definition,  
## but +towgs84= values preserved



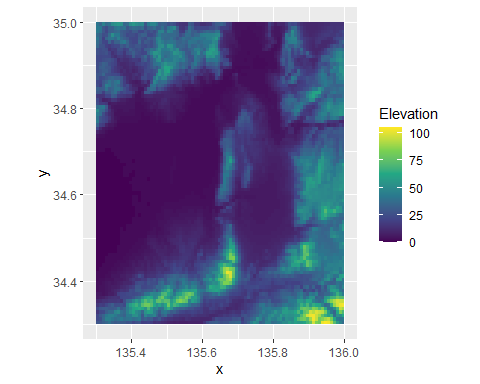
summary(crop\_DEM) #切り出したラスターデータの概要を表示

## el  
## Min. 0  
## 1st Qu. 4  
## Median 14  
## 3rd Qu. 33  
## Max. 105  
## NA's 0

We can plot the raster data using ggplot function, which allows us to modify axis, legend, and labels displayed on the plot. To use ggplot, we need to convert the cropped raster to a dataframe.

ggplot関数を使用して、地図上に表示される軸や凡例、ラベルを編集し、ラスターデータを描画することができます。そのために、切り出したラスターデータをデータフレームに変換する必要があります。

# cover to a dataframe for ggplot／データフレームへの変換 ※ggplot用  
crop\_DEM\_df <- as.data.frame(crop\_DEM, xy = TRUE)  
# plot  
ggplot() +  
 geom\_raster(data = crop\_DEM\_df , aes(x = x, y = y, fill = el)) +  
 scale\_fill\_viridis\_c(name = "Elevation") +  
 coord\_quickmap() # plot faster／描画速度向上



Now, let’s work on vector data and plot it on the raster. We import a shapefile which contains site locations we want to explore (note: its not a real data). We are curious about the distribution of archaeological sites and how they relate to the elevation of this area.

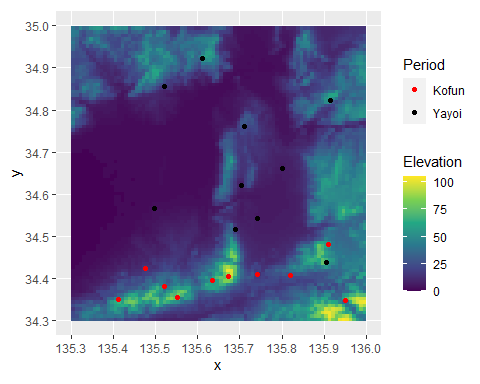
続いて、ラスターデータの上にベクターデータを描画しましょう。検討したい遺跡の位置情報(ダミー：実在のデータではありません)を含むシェープファイルを読み込みます。この地域における標高と遺跡分布の関係に関心があります。

## Exercise 5: explore shapefile and map it on the raster layer (7 mins)／実習5：ラスターデータ・レイヤーでshapeファイルと地図を検討

crop\_DEM\_df <- as.data.frame(crop\_DEM, xy = TRUE)  
# Example of archaeological sites／考古遺跡データのサンプル   
sites\_location <- st\_read("data/workshop\_data/sites\_example.shp") #shpファイル読込

## Reading layer `sites\_example' from data source `C:\Users\Atsushi Noguchi\Documents\GitHub\Japan\_GIS\_workshop\_202006\_LW\data\workshop\_data\sites\_example.shp' using driver `ESRI Shapefile'  
## Simple feature collection with 20 features and 2 fields  
## geometry type: POINT  
## dimension: XY  
## bbox: xmin: 135.412 ymin: 34.349 xmax: 135.951 ymax: 34.923  
## CRS: NA

crop\_DEM\_df %>%   
 ggplot() +   
 geom\_raster(aes(x = x, y = y, fill = el)) +  
 geom\_sf(data = sites\_location, aes(color = Period)) + # add site shapefile／遺跡のシェープを追加   
 scale\_color\_manual(values=c("red", "black")) + # change default color／基本色を変更   
 scale\_fill\_viridis\_c(name = "Elevation") +  
 coord\_sf() # all layers use a common CRS／共通のCRS(参照座標系)をすべてのレイヤーに使用

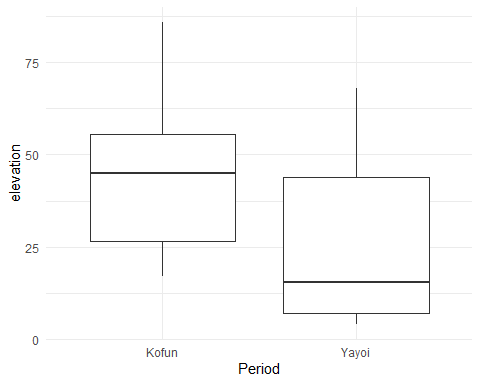


We are curious about the elevation of the locations of archaeological sites, and would like to compare sites from two phases: Yayoi and Kofun.

弥生時代と古墳時代、2つの時期の遺跡の立地・標高を比較してみましょう。

## Exercise 6: extract elevation and make a plot to compare sites from two phases (5 mins)／各遺跡の標高値を取得し2つの時代を比較するグラフを描画

# convert sf (simple feature) to a spatial object／空間データ・オブジェクトをsf(simple feature)に変換   
sp\_sites\_location <- as(sites\_location, "Spatial")  
  
# extract elevation for each site／各遺跡の標高を取得  
elevation <- extract(crop\_DEM, sp\_sites\_location,   
 method = "simple") # use values for the cell a point falls in／遺跡が描画される位置のメッシュの情報をを使用   
sites\_location <- cbind(sites\_location, elevation) #取得した標高値をデータ・オブジェクトに追加   
  
sites\_location %>% #2つの時代の遺跡の標高の箱ひげ図を描画   
 ggplot(aes(Period, elevation)) +   
 geom\_boxplot() +  
 theme\_minimal()



# Density analysis and hypothesis testing／分布密度の分析と仮説の検証

We may want to know the distribution pattern of the sites across this area. We can visualize the density to check any hot spots using kernal density estimation.

次に、この地域の遺跡分布のパターンを検討したいと思います。カーネル密度推定により遺跡分布密度を可視化し「ホットスポット」があるかどうかを確認します。

## Exercise 7: make a kernel density plot (7 mins)／実習7：カーネル密度図の描画

library(spatstat)

## Loading required package: spatstat.data

## Loading required package: nlme

##   
## Attaching package: 'nlme'

## The following object is masked from 'package:raster':  
##   
## getData

## The following object is masked from 'package:dplyr':  
##   
## collapse

## Loading required package: rpart

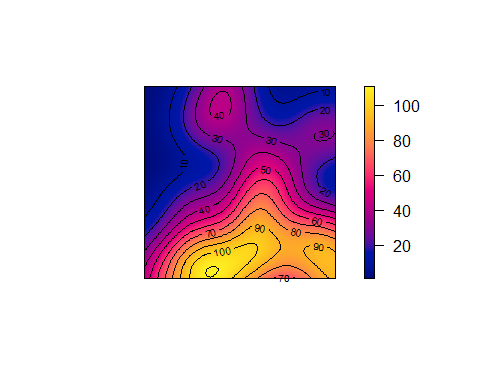
## Registered S3 method overwritten by 'spatstat':  
## method from  
## print.boxx cli

##   
## spatstat 1.64-1 (nickname: 'Help you I can, yes!')   
## For an introduction to spatstat, type 'beginner'

##   
## Attaching package: 'spatstat'

## The following objects are masked from 'package:raster':  
##   
## area, rotate, shift

library(maptools)  
crop\_DEM\_df <- as.data.frame(crop\_DEM, xy = TRUE)  
  
# get two columns, one longitude and another is latitude／経度と緯度の取得   
sites\_location\_coords <-  
 sites\_location %>%   
 st\_coordinates() %>%   
 as.data.frame   
  
# convert to ppp object that represent a two-dimensional point pattern／2次元の点分布パターンを反映するpppオブジェクトを変換  
sites\_location\_ppp <- ppp(x = sites\_location\_coords$X,  
 y = sites\_location\_coords$Y,  
 range(crop\_DEM\_df$x), # set window, means the extent of an area  
 range(crop\_DEM\_df$y))  
  
K1 <- density(sites\_location\_ppp)   
  
plot(K1, main=NULL, las=1)  
contour(K1, add=TRUE)

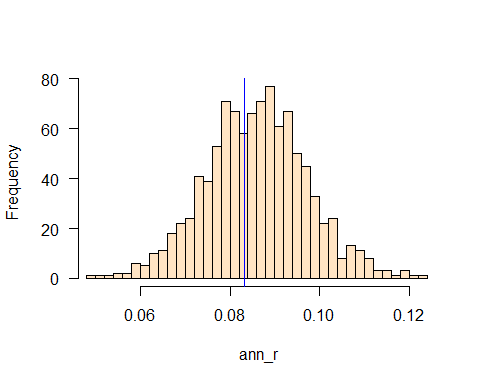


Is the hot spots we observed significant? We can simulate the site locations and testing our hypothesis to determine if the distribution is random or not random.

明確な「ホットスポット」は確認できましたか? 次に遺跡位置をシミュレートし、分布パターンがランダムかどうかを検証します。

## Exercise 8: simulation and plot the histogram (7 mins)／実習8：シミュレーションとヒストグラムの描画

# get the mean distance for our observation／観察結果から平均距離を取得  
ann\_p <- mean(nndist(sites\_location\_ppp, k=1))  
n <- 1000 # Number of simulations／シミュレーションの試行回数   
  
ann\_r <- vector(length = n) # an object for storing simulated ANN values／シミュレーションしたANNをann\_rに収納   
  
# simulation／シミュレーション  
for (i in 1:n){ #試行回数を指定(n=1000)  
 rand\_p <- rpoint(sites\_location\_ppp$n,   
 win = as.owin(crop\_DEM\_df)) # generate random point locations／ランダムな位置を生成  
 ann\_r[i] <- mean(nndist(rand\_p, k=1)) # tally the ANN values／ANNを集計  
}  
  
# plot the histogram and add our observed ANN value line／ヒストグラムを描画し、シミュレーション結果をのラインを追加  
hist(ann\_r, main=NULL, las=1, breaks=40,   
 col = "bisque",   
 xlim = range(ann\_p, ann\_r))  
abline(v = ann\_p, col="blue") # the observed value

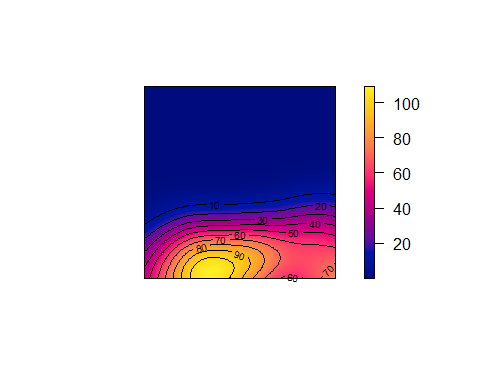


We have looked at the distribution of sites all together, but what if we want to focus on sites from a phase; for example, we want to explore Kofun period. We can filter out the phase we want and then use the same method to test the Kofun sites.

ここでは全ての遺跡を検討しましたが、たとえば古墳時代の遺跡分布を検証するためには、遺跡分布データを時代でフィルターして同じ方法を適用します。

## Quiz: Do the Kofun sites randomly distributed or non-randomly distributed? The pattern is clustterd or dispersed? ／クイズ：古墳時代遺跡はランダムな分布ですか? 分布パターンはクラスター状ですか? それとも分散分布ですか?

sites\_location\_coords\_kofun <- #古墳時代遺跡のカーネル密度推定と描画   
 sites\_location %>%   
 st\_coordinates() %>%   
 as.data.frame () %>%   
 bind\_cols(sites\_location) %>%   
 filter(Period == "Kofun")  
  
sites\_location\_ppp\_kofun <- (ppp(x = sites\_location\_coords\_kofun$X,  
 y = sites\_location\_coords\_kofun$Y,  
 range(crop\_DEM\_df$x),  
 range(crop\_DEM\_df$y)))   
  
K2 <- density(sites\_location\_ppp\_kofun)   
  
plot(K2, main=NULL, las=1)  
contour(K2, add=TRUE)



ann\_p <- mean(nndist(sites\_location\_ppp\_kofun, k=1))  
n <- 1000 # Number of simulations  
  
ann\_r <- vector(length = n) # an object for storing simulated ANN values  
  
# simulation  
for (i in 1:n){  
 rand\_p <- rpoint(sites\_location\_ppp\_kofun$n,   
 win = as.owin(crop\_DEM\_df)) # Generate random point locations  
 ann\_r[i] <- mean(nndist(rand\_p, k=1)) # Tally the ANN values  
}  
  
# plot the histogram and add our observed ANN value line  
hist(ann\_r, main=NULL, las=1, breaks=40,   
 col = "bisque",   
 xlim = range(ann\_p, ann\_r))  
abline(v = ann\_p, col="blue")

