

INTRODUCCIÓN A LA GEOCOMPUTACIÓN CON R

EXAMEN FINAL

FECHA DE ENTREGA: DICIEMBRE 2, 2022

1. Considera el siguiente código

```
f <- function(x){  
  x^2  
}
```

y responde

- (a) (1 punto) ¿Cuál es el resultado de la operación $f(1) + f(f(2)) * f(f(f(3)))$?
- (b) (1 punto) Grafica la función $f(x)+2*x$ cuando x toma valores en el vector `seq(-5,5,by=1)`.
- (c) (1 punto) ¿En qué punto del vector `seq(-5,5,by=1)`, la función $f(x)+2*x$ alcanza su valor mínimo? Justifica tu respuesta.
2. El operador `%*%` permite multiplicar dos matrices. En este ejercicio definirás una función alternativa al operador `%*%`. Iniciemos con el caso particular: Considera las matrices

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}, \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix}.$$

(a) (1 punto) Multiplicar el primer renglón de la matriz A con la primera columna de la matriz B consiste en dos operaciones: (i) multiplicar (entrada a entrada) estos dos vectores, y (ii) sumar los valores del vector resultante en (i). En símbolos:

$$(1) \quad C[1, 1] = A[1,] \times B[, 1] = \sum_{k=1}^3 A[1, k] \times B[k, 1].$$

Implementa la operación de la ecuación (1).

(b) (0.75 puntos) La operación $A \times B$ consiste en multiplicar cada renglón de A con cada columna de B utilizando el mismo principio presentado en (a). Es decir,

$$(2) \quad C[i, j] = A[i,] \times B[, j] = \sum_{k=1}^3 A[i, k] \times B[k, j], \quad i = 1, 2, 3; j = 1, 2, 3.$$

Implementa la operación de la ecuación (2). **Nota:** El resultado de esta operación debe ser una matrix, llámala C , con 3 renglones y 3 columnas.

Ahora resolvamos el caso general. Considera las matrices

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{n,m} \end{pmatrix}, \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,q} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,q} \end{pmatrix}.$$

(c) (1 punto) Implementa la operación

$$(3) \quad C[i, j] = A[i, :] \times B[:, j] = \sum_{k=1}^m A[i, k] \times B[k, j], \quad i = 1, \dots, n; j = 1, \dots, q.$$

(d) (1.25 puntos) Escribe una función cuyos argumentos sean 2 matrices arbitrarias (A y B) y la salida sea el producto $A \times B$. **Nota:** La función debe verificar que el número de columnas de A sea igual al número de renglones de B ; la salida de la función debe ser una matriz con número de renglones y columnas igual al número de renglones de A y columnas de B , respectivamente.

3. En la carpeta `examenFinal` encontrarás el archivo `Investigadores-SNI-Vigentes-2018.csv` el cual contiene la base de datos de los miembros del Sistema Nacional de Investigadores (SNI) vigentes en el año 2018.

(a) (0.5 puntos) ¿Cuáles son las variables de esta base de datos? ¿Cuántos investigadores estaban vigentes en el SNI en 2018?

(b) (2 puntos) ¿Cuál es el área de conocimiento con más investigadoras e investigadores? De esta área, ¿cuáles son las 5 instituciones con más investigadores adscritos?

(c) (1 punto) En 2018 ¿cuántos investigadores de la UNAM pertenecían al SNI? De este grupo, ¿cuántos investigadores pertenecían al Área I? Finalmente, de este subgrupo ¿cuántos pertenecían al Nivel 3?

(d) (0.5 puntos) Considera la sub-base de datos correspondiente a los investigadores de la UNAM que en 2018 pertenecían a Área I en el Nivel 3 del SNI. Ordena alfabéticamente esta sub-base de datos. **Nota:** es suficiente ordenar por apellido paterno, por ejemplo, las primeras 5 entradas de la sub-base de datos *ya* ordenada luce así:

```
##          Grado Apellido.Paterno Apellido.Materno      Nombre Nivel
## 24775    DR.          ACOSTA          NAJARRO DWIGHT ROBERTO      3
## 25747   DRA.          ALANIZ          ALVAREZ  SUSANA ALICIA      3
## 25247    DR.          ALBA           ANDRADE   FERNANDO      3
## 25147    DR.      ALCUBIERRE          MOYA     MIGUEL      3
## 24718    DR.          ALONSO          NUÑEZ    GABRIEL      3
##
##          Institución.de.Adscripción
## 24775 UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
## 25747 UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
## 25247 UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
## 25147 UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
## 24718 UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
##
##          Área.del.Conocimiento
## 24775 Área I: FISICO-MATEMÁTICAS Y CIENCIAS DE LA TIERRA
## 25747 Área I: FISICO-MATEMÁTICAS Y CIENCIAS DE LA TIERRA
## 25247 Área I: FISICO-MATEMÁTICAS Y CIENCIAS DE LA TIERRA
## 25147 Área I: FISICO-MATEMÁTICAS Y CIENCIAS DE LA TIERRA
## 24718 Área I: FISICO-MATEMÁTICAS Y CIENCIAS DE LA TIERRA
```

4. En esta clase te has familiarizado con el uso de los siguientes paquetes de R:

```
doParallel,foreach,geoTS,gtools,imputeTS,mapview,raster,RColorBrewer,rgdal,sf
```

En este ejercicio vas a explorar la popularidad de estos paquetes a través de registrar sus descargas diarias. Para este fin necesitarás el paquete `cranlogs` el cual puedes descargar desde GitHub con la instrucción

```
devtools::install_github("metacran/cranlogs")
```

Para mayores detalles puedes también mirar esta [liga](#).

Haciendo uso de este paquete responde:

- (1 punto) ¿Cuántas descargas tuvieron en el **último mes** cada uno de los 10 paquetes que hemos utilizado en este seminario?
- (1 punto) ¿Cuántas descargas **totales** tuvieron desde su fecha de publicación hasta hoy cada uno de los 10 paquetes que hemos utilizado en este seminario? Para simplificar el problema, para cada paquete considera únicamente la versión que tienes instalada en tu sistema.
- (1 punto) De acuerdo con el inciso anterior, ¿cuál es el **promedio mensual** de descargas por cada paquete?
- (1 punto) Haz un `data.frame` con tus respuestas a los incisos anteriores. Por ejemplo, para el paquete `raster` tu solución podría parecerse a esto:

```
## Paquete Descargas.Ultimo.Mes Descargas.Totales Descargas.Promedio.Mensual
## 7 raster 177695 839847 167969
```

5. En este ejercicio harás uso de los paquetes `sf`, `spData` y `tmap` por lo que debes asegurarte de tener estos paquetes, y todas sus dependencias, instalados en tu sistema.

Los datasets `nz` and `nz_height` contienen información sobre las 16 regiones principales de Nueva Zelanda y sus 101 picos más altos, respectivamente:

```
str(nz)
```

```
## Classes 'sf' and 'data.frame': 16 obs. of 7 variables:
## $ Name : chr "Northland" "Auckland" "Waikato" "Bay of Plenty" ...
## $ Island : chr "North" "North" "North" "North" ...
## $ Land_area : num 12501 4942 23900 12071 8386 ...
## $ Population : num 175500 1657200 460100 299900 48500 ...
## $ Median_income: int 23400 29600 27900 26200 24400 26100 29100 25000 32700 26900 ...
## $ Sex_ratio : num 0.942 0.944 0.952 0.928 0.935 ...
## $ geom :sfc_MULTIPOLYGON of length 16; first list element: List of 1
## ..$ :List of 1
## ...$ : num [1:68, 1:2] 1745493 1740539 1733165 1720197 1709110 ...
## ..- attr(*, "class")= chr [1:3] "XY" "MULTIPOLYGON" "sfg"
## - attr(*, "sf_column")= chr "geom"
## - attr(*, "agr")= Factor w/ 3 levels "constant","aggregate",...: NA NA NA NA NA NA
## ..- attr(*, "names")= chr [1:6] "Name" "Island" "Land_area" "Population" ...
```

```
str(nz_height)
```

```
## Classes 'sf' and 'data.frame': 101 obs. of 3 variables:
## $ t50_fid : int 2353944 2354404 2354405 2369113 2362630 2362814 2362817 2363991 2363993 2
## $ elevation: int 2723 2820 2830 3033 2749 2822 2778 3004 3114 2882 ...
```

```
## $ geometry :sfc_POINT of length 101; first list element: 'XY' num 1204143 5049971
## - attr(*, "sf_column")= chr "geometry"
## - attr(*, "agr")= Factor w/ 3 levels "constant","aggregate",...: NA NA
## ..- attr(*, "names")= chr [1:2] "t50_fid" "elevation"
```

En particular, con 70, Canterbury es la región con la mayor cantidad de picos en Nueva Zelanda, véase la Figura 1:

```
canterbury <- nz[nz$Name=="Canterbury",]
canterbury_height <- nz_height[canterbury,]

plot(nz["geom"], reset=FALSE)
plot(canterbury["Name"], col="gray", add=TRUE)
plot(canterbury_height["elevation"], col="red", add=TRUE)
```



FIGURA 1. Mapa de la región Canterbury, Nueva Zelanda (en gris) junto con los picos más altos en la región.

```
str(canterbury_height)
```

```
## Classes 'sf' and 'data.frame': 70 obs. of 3 variables:
## $ t50_fid : int 2362630 2362814 2362817 2363991 2363993 2363994 2363995 2363997 2363998 2
## $ elevation: int 2749 2822 2778 3004 3114 2882 2796 3070 3061 3077 ...
## $ geometry :sfc_POINT of length 70; first list element: 'XY' num 1378170 5158491
## - attr(*, "sf_column")= chr "geometry"
## - attr(*, "agr")= Factor w/ 3 levels "constant","aggregate",...: NA NA
```

```
## ..- attr(*, "names")= chr [1:2] "t50_fid" "elevation"
```

(a) (1.5 puntos) ¿Cuál es la región con la segunda mayor cantidad de picos en Nueva Zelanda?

(b) (1.5 puntos) De acuerdo a estos datasets ¿hay regiones en Nueva Zelanda sin picos? ¿Cuáles son estas regiones?

(c) (2 puntos) Utiliza la función `aggregate()` para producir un mapa estático de elevación promedio en cada región de Nueva Zelanda. **Nota:** Véase Figura 2.

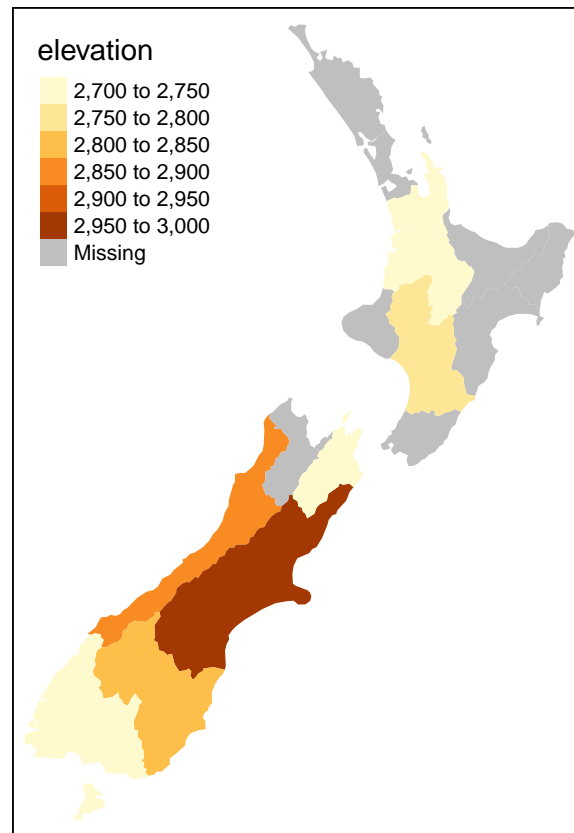


FIGURA 2. Elevación promedio de regiones en Nueva Zelanda.