

Presentacion metodos

GG3

2024-08-14

Precipitacion Mensual en Cundinamarca

Por: Juan Camilo Aguirre

Metodos Estadisticos

Agosto 14, 2024

Introduccion

En esta presentacion, se mostrara el analisis comprensivo de la base de datos: *Precipitaciones Totales Mensuales*.

Volumen de precipitaciones mensuales obtenido en la red de estaciones hidrológicas de la Corporación Autónoma Regional de Cundinamarca - CAR.

Esta trae informacion sobre precipitaciones en el departamento de Cundinamarca desde 1920 hasta 2019.

Las dimensiones son de 6,196 (filas) x 19 (columnas) = 117,724 datos.

Objetivos de la presentacion

(1. ¿Qué objetivo plantearías para el estudio o el interés del investigador?)

- Realizar un analisis descriptivo sobre los datos climatologicos.
- Estalecer tendencias sobre los diversos meses y su respectiva precipitacion.

Para la presentacion, se utilizaron las siguientes librerias como herramientas para manipular y organizar los datos:

```
## Warning: package 'readxl' was built under R version 4.3.3
```

```
## Loading required package: viridisLite
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
## Warning: package 'sf' was built under R version 4.3.3
```

```
## Linking to GEOS 3.11.2, GDAL 3.8.2, PROJ 9.3.1; sf_use_s2() is TRUE
```

```
## Warning: package 'ggspatial' was built under R version 4.3.3
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

## Warning: package 'pastecs' was built under R version 4.3.3

##
## Attaching package: 'pastecs'

## The following objects are masked from 'package:dplyr':
##
##   first, last

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

## Warning: package 'patchwork' was built under R version 4.3.3

## Warning: package 'leaflet' was built under R version 4.3.3

## Warning: package 'summarytools' was built under R version 4.3.3

## Warning: package 'pandoc' was built under R version 4.3.3
```

Con esto en mente, podemos proceder con la siguiente parte de la presentacion.

Cuadro operacional de variables

(2.¿Qué variables se involucran en el estudio y de qué tipo son? Cuadro de operacionalización de variables)

Variables	Tipo de variable	Descripción	Unidad de medida	Valores	Operacionalización
ESTACION	Cualitativa nominal	Estación donde se recolectaron los datos.	Texto	Nombres de estaciones (Pajas blancas, etc.)	Registro del nombre de las estaciones
CODIGO	Cualitativa ordinal	Código numérico de la estación	Numero entero	Valores con números (códigos)	Registro del código de la estación.
X=N e Y=E	Cuantitativa Continua	Coordenadas UTM en los ejes X e Y.	Metros	Valores de coordenadas.	Medición horizontal y vertical de la posición de la estación.
MUNICIPIO	Cualitativa nominal	Nombre del municipio de la estación.	Texto	Nombres de municipios	Registro del nombre del municipio
CUENCA	Cualitativa nominal	Nombre de la cuenca de la estación.	Texto	Nombres de cuencas	Registro del nombre de la cuenca.
AÑO	Cuantitativa discreta	Año donde se recolectaron los datos	Numero entero (año)	Años	Años donde se recolectaron la precipitación.
MES (ENERO, FEBRERO, etc)	Cuantitativa continua	Meses donde se recolectaron los datos	Milímetros (mm)	Valores numéricos de precipitación.	Precipitación acumulada en los meses del año.

Pre-Análisis descriptivo

```
#Primero se carga la base de datos
precipitaciones <- read_excel("C:\\Users\\Administrador\\Documents\\R\\precipitaciones\\precipitaciones
head(precipitaciones)
```

```
## # A tibble: 6 x 19
##   ESTACIÓN CÓDIGO 'X=N' 'Y=E' MUNICIPIO CUENCA AÑO ENERO FEBRERO MARZO ABRIL
##   <chr> <dbl> <dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 1995 0 0 0 0
## 2 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 1996 296. 127. 248. 149.
## 3 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 1997 294. 76.9 160. 296.
## 4 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 1998 144. 118 209. 438.
## 5 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 1999 78.3 257. 83 339.
## 6 Pajas B~ 2.12e6 973320 944710 Nilo R. Pa~ 2000 113. 305. 202. 323.
## # i 8 more variables: MAYO <dbl>, JUNIO <dbl>, JULIO <dbl>, AGOSTO <dbl>,
## # SEPTIEMBRE <dbl>, OCTUBRE <dbl>, NOVIEMBRE <dbl>, DICIEMBRE <dbl>
```

```
hist(precipitaciones$AÑO, breaks=75,
     main="Histograma de Precipitaciones",
     xlab="Año",
     ylab="Frecuencia",
     sub="Figura 1. Histograma de frecuencias de registros de precipitaciones.")
```

Histograma de Precipitaciones

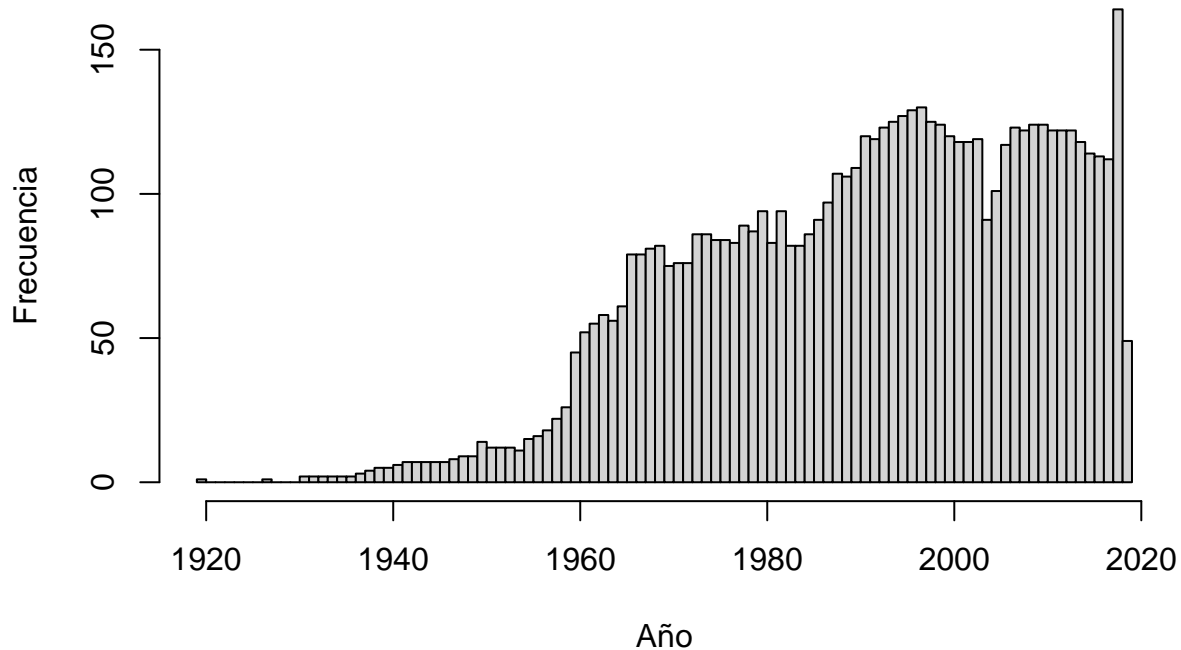


Figura 1. Histograma de frecuencias de registros de precipitaciones.

Para encontrar resultados numericos enfocados en los datos importantes, creamos un nuevo dataset con solo los meses:

```
meses <- precipitaciones %>% select(ENERO:DICIEMBRE)
```

Aqui se sacaran dos diferentes metodos para obtener resultados numericos:

Metodo 1: `summary()`

```
summary(meses)
```

```
##      ENERO      FEBRERO      MARZO      ABRIL
## Min.   : 0.00  Min.   : 0.00  Min.   : 0.00  Min.   : 0.00
## 1st Qu.: 8.60  1st Qu.: 19.18  1st Qu.: 37.77  1st Qu.: 59.30
## Median : 26.05  Median : 40.45  Median : 71.40  Median : 98.75
## Mean   : 39.42  Mean   : 52.69  Mean   : 83.61  Mean   :115.18
## 3rd Qu.: 53.40  3rd Qu.: 71.00  3rd Qu.: 113.40  3rd Qu.:154.50
## Max.   :3097.00  Max.   :2797.00  Max.   :1298.00  Max.   : 828.50
##
##      MAYO      JUNIO      JULIO      AGOSTO
## Min.   : 0.0  Min.   : 0.00  Min.   : 0.00  Min.   : 0.00
## 1st Qu.: 53.5  1st Qu.: 29.70  1st Qu.: 23.00  1st Qu.: 22.20
## Median : 90.2  Median : 55.50  Median : 44.15  Median : 42.15
## Mean   :100.4  Mean   : 64.86  Mean   : 55.39  Mean   : 51.69
```

```
## 3rd Qu.:134.6 3rd Qu.: 88.00 3rd Qu.: 72.70 3rd Qu.: 69.50
## Max. :823.5 Max. :791.00 Max. :550.00 Max. :500.50
## NA's :1
## SEPTIEMBRE OCTUBRE NOVIEMBRE DICIEMBRE
## Min. : 0.00 Min. : 0.0 Min. : 0.0 Min. : 0.00
## 1st Qu.: 28.07 1st Qu.: 58.5 1st Qu.: 52.9 1st Qu.: 18.20
## Median : 52.00 Median :102.0 Median : 91.5 Median : 44.05
## Mean : 61.79 Mean :113.5 Mean :105.3 Mean : 57.10
## 3rd Qu.: 83.72 3rd Qu.:153.3 3rd Qu.:140.5 3rd Qu.: 80.53
## Max. :498.60 Max. :675.2 Max. :767.0 Max. :532.30
##
```

Metodo 2: stat.desc() (usando la libreria pastecs)

```
stat.desc(meses)
```

```
##          ENERO      FEBRERO      MARZO      ABRIL      MAYO
## nbr.val    6.196000e+03 6.196000e+03 6.196000e+03 6.196000e+03 6.196000e+03
## nbr.null    6.270000e+02 3.920000e+02 3.280000e+02 3.000000e+02 3.560000e+02
## nbr.na      0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## min        0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## max        3.097000e+03 2.797000e+03 1.298000e+03 8.285000e+02 8.235000e+02
## range      3.097000e+03 2.797000e+03 1.298000e+03 8.285000e+02 8.235000e+02
## sum        2.442283e+05 3.264612e+05 5.180772e+05 7.136501e+05 6.218304e+05
## median     2.605000e+01 4.045000e+01 7.140000e+01 9.875000e+01 9.020000e+01
## mean       3.941709e+01 5.268903e+01 8.361478e+01 1.151792e+02 1.003600e+02
## SE.mean    8.858229e-01 7.626290e-01 8.549761e-01 1.052163e+00 8.876913e-01
## CI.mean.0.95 1.736520e+00 1.495018e+00 1.676050e+00 2.062605e+00 1.740183e+00
## var        4.861891e+03 3.603612e+03 4.529178e+03 6.859266e+03 4.882423e+03
## std.dev    6.972726e+01 6.003010e+01 6.729917e+01 8.282069e+01 6.987433e+01
## coef.var   1.768960e+00 1.139328e+00 8.048716e-01 7.190596e-01 6.962371e-01
##          JUNIO      JULIO      AGOSTO      SEPTIEMBRE      OCTUBRE
## nbr.val    6.195000e+03 6.196000e+03 6.196000e+03 6.196000e+03 6.196000e+03
## nbr.null    4.590000e+02 4.840000e+02 4.780000e+02 4.280000e+02 4.060000e+02
## nbr.na      1.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## min        0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## max        7.910000e+02 5.500000e+02 5.005000e+02 4.986000e+02 6.752000e+02
## range      7.910000e+02 5.500000e+02 5.005000e+02 4.986000e+02 6.752000e+02
## sum        4.018006e+05 3.432026e+05 3.202993e+05 3.828326e+05 7.032189e+05
## median     5.550000e+01 4.415000e+01 4.215000e+01 5.200000e+01 1.020000e+02
## mean       6.485885e+01 5.539099e+01 5.169453e+01 6.178706e+01 1.134956e+02
## SE.mean    6.711454e-01 6.413230e-01 5.747515e-01 6.350710e-01 1.018800e+00
## CI.mean.0.95 1.315678e+00 1.257216e+00 1.126712e+00 1.244960e+00 1.997202e+00
## var        2.790452e+03 2.548385e+03 2.046782e+03 2.498941e+03 6.431165e+03
## std.dev    5.282473e+01 5.048153e+01 4.524138e+01 4.998941e+01 8.019455e+01
## coef.var   8.144567e-01 9.113671e-01 8.751677e-01 8.090596e-01 7.065871e-01
##          NOVIEMBRE      DICIEMBRE
## nbr.val    6.196000e+03 6.196000e+03
## nbr.null    4.430000e+02 5.690000e+02
## nbr.na      0.000000e+00 0.000000e+00
## min        0.000000e+00 0.000000e+00
## max        7.670000e+02 5.323000e+02
```

```
## range      7.670000e+02 5.323000e+02
## sum        6.525248e+05 3.537983e+05
## median     9.150000e+01 4.405000e+01
## mean       1.053139e+02 5.710108e+01
## SE.mean    1.015103e+00 6.964416e-01
## CI.mean.0.95 1.989954e+00 1.365267e+00
## var        6.384568e+03 3.005252e+03
## std.dev    7.990349e+01 5.482017e+01
## coef.var   7.587176e-01 9.600549e-01
```

Adicionalmente, se filtraran la cantidad de registros por estacion.

```
station_count <- precipitaciones %>%
  group_by(ESTACIÓN) %>%
  summarize(Conteo = n(), .groups = 'drop') %>%
  arrange(ESTACIÓN)

print(station_count)
```

```
## # A tibble: 174 x 2
##   ESTACIÓN      Conteo
##   <chr>         <int>
## 1 Acandy         42
## 2 Agua Fría      22
## 3 Alco           37
## 4 Almaviva       21
## 5 Alto Caicedo    4
## 6 Alto La Puerta  3
## 7 Alto Saboya     36
## 8 Alto de Aire    33
## 9 América        10
## 10 Anclaje 14     31
## # i 164 more rows
```