**Vargas Carrillo Raul Paulino**

**Correlación**

Obtener la información utilizando R deberá adjuntar tanto el código utilizado como los resultados obtenidos.

Para los conjuntos de datos especificados en cada uno de los retos, será necesario cal cular la Covarianza, el Coeficiente de Correlación y crear una gráfico de dispersión.

1. **Reto**

El archivo de Excel precipitación contienen la información de una extensa región de pastizales donde se ubicaron 4 áreas con diferentes valores de Precipitación pluvial anual conocidos de antemano (respectivamente 300 mm,600 mm, 900 mm y 1200 mm). Dentro de cada área se eligieron al azar 4 parcelas de 100 m² cada una y en cada parcela medir la productividad primaria neta del pastizal (PPN).

|  |  |  |  |
| --- | --- | --- | --- |
| Precipitación | ppn | Precipitación | ppn |
| 300 | 80 | 900 | 560 |
| 300 | 150 | 900 | 590 |
| 300 | 125 | 900 | 400 |
| 300 | 170 | 900 | 380 |
| 600 | 390 | 1200 | 790 |
| 600 | 470 | 1200 | 610 |
| 600 | 225 | 1200 | 640 |
| 600 | 350 | 1200 | 680 |

> precipitacion %>%

+ ggplot(aes(Precipitacion, ppn)) +

+ geom\_point()

> cov(precipitacion$Precipitacion, precipitacion$ppn)

[1] 70800

> cor(precipitacion$Precipitacion, precipitacion$ppn)

[1] 0.9308741

La correlación de las variables es de tipo directa o correlación positiva, además, es una correlación fuerte.

Gráfico

Descripción generada automáticamente

**2. Reto**

El archivo de Exel de ventas contiene la información de una muestra de 15 super mercados para las variables número de empleados y ventas semanales en cientos de miles de pesos.

|  |  |  |  |
| --- | --- | --- | --- |
| NumEmpleado | Ventas | NumEmpleado | Ventas |
| 10 | 4 | 32 | 11 |
| 17 | 6 | 32 | 13 |
| 17 | 6 | 36 | 15 |
| 17 | 5 | 40 | 16 |
| 19 | 9 | 41 | 16 |
| 24 | 8 | 43 | 17 |
| 24 | 12 | 48 | 20 |
| 30 | 12 |  |  |

> ventas %>%

+ ggplot(aes(NumEmpleado, Ventas)) +

+ geom\_point()

> cov(ventas$NumEmpleado, ventas$Ventas)

[1] 54.40476

> cor(ventas$NumEmpleado, ventas$Ventas)

[1] 0.9719477

Imagen que contiene exterior, foto, agua, gente

Descripción generada automáticamenteAl igual que con los datos del reto anterior, la correlación de las variables es de tipo directa o positiva, además, es una correlación fuerte.

**3. Reto**

Tiempo de espera entre erupciones y duración de la erupción del geiser Old Faithful en el Parque Nacional Yellowstone, Wyoming, EE. UU. (faithful.xlsx)

eruptions: Tiempo de erupción en minutos

waiting: Tiempo de espera hasta la próxima erupción (en minutos)

|  |  |
| --- | --- |
| eruptions | waiting |
| 3.6 | 79 |
| 1.8 | 54 |
| 3.333 | 74 |
| 2.283 | 62 |
| 4.533 | 85 |
| 2.883 | 55 |
| 4.7 | 88 |
| 3.6 | 85 |
| 1.95 | 51 |
| 4.35 | 85 |
| 1.833 | 54 |
| 3.917 | 84 |
| 4.2 | 78 |
| 1.75 | 47 |
| 4.7 | 83 |
| 2.167 | 52 |
| 1.75 | 62 |
| 4.8 | 84 |
| 1.6 | 52 |

> faithful %>%

+ ggplot(aes(eruptions, waiting)) +

+ geom\_point()

> cov(faithful$eruptions, faithful$waiting)

[1] 13.97781

> cor(faithful$eruptions, faithful$waiting)

[1] 0.9008112

Gráfico, Gráfico de dispersión

Descripción generada automáticamenteLas variables tienen una correlación positiva, con una correlación fuerte.