



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE GOBIERNO
GOB3032 - ANÁLISIS DE DATOS I

Tarea 1

Análisis de Datos I

05 de abril 2024

Instrucciones

Se espera que entregue sus respuestas en un archivo pdf en conjunto con el código (script) utilizado en el buzón de canvas designado hasta el día **11 de abril a las 23:59**.

Pregunta 1

La siguiente tabla contiene información ficticia de la calidad de vida en distintas ciudades del mundo.

Ciudad	Tamaño	Ingreso (USD)	Índice de Criminalidad	Calidad del Aire	Nivel de Desarrollo
París	Grande	60000	65	6.8	Alto
Zurich	Mediana	75000	30	8.0	Alto
Estambul	Grande	38000	58	6.7	Medio
Montevideo	Pequeña	40000	30	7.0	Bajo
Barcelona	Mediana	58000	35	7.5	Medio
Oslo	Mediana	72000	25	7.8	Alto
Londres	Grande	58000	30	7.0	Alto
Santiago	Grande	45000	38	6.3	Medio

- En base a los datos de la tabla genere un vector para cada una de las columnas y luego genere un data-frame con las 6 columnas creadas.
- Reporte las medidas de tendencia central y dispersión de las variables correspondientes.
- Para las variables cualitativas presente una tabla de frecuencia relativa.

Pregunta 2

Del mismo modo que utilizamos **ggplot2** para graficar histogramas, podemos generar gráficos de dispersión para visualizar la relación entre dos variables x e y en una base de datos `df` con el siguiente código.

```
library(ggplot2)
ggplot(df, aes(x=x, y=y)) + geom_point()
```

Además de visualizar esta relación, podemos describirla en base a su correlación, que indica la fuerza y la dirección de esta. En `r` podemos estimar la correlación entre dos variables con el siguiente código.

```
cor(x,y)
```

Considerando lo anterior responda las siguientes preguntas.

- a. Haga un grafico de dispersión con la variables de ingreso y criminalidad. Estime el coeficiente de correlación e interprete los resultados.
- b. Haga un grafico de dispersión con la variables de ingreso y calidad del aire. Estime el coeficiente de correlación e interprete los resultados.