Solución

1.1) La única variable categórica es "city", entonces decido no describirla ya que no es relevante para el análisis, porque sus datos no se repiten, y la única relación que tienen entre sí es que son ciudades de Colombia

Pero le voy a calcular esto a una variable numérica (cuantitativa) de la tabla. (GDP (USD billion))

Medíana

→ Ordeno

0.6 ; 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.5, 1.7, 1.8, 2, 2.1, 2.3, 2.5, 2.8, 3.0, 3.2, 3.5, 3.8, 4, 4.8, 5.1, 6.2, 7.3, 10.5, 16.8, 22.4, 44.1, 103.5.

→ Tengo 30 datos

$$\Rightarrow \frac{2{,}8 + 2{,}5}{2} = 2{,}65$$

$$\text{Media} = \frac{\sum \text{datos}}{N}$$

$$\Rightarrow \frac{262{,}5}{30} = 8{,}75$$

→ Moda = Ningun dato se repite así que ~~no~~ escribiré todos

→ desviación estandair

Como la tabla no dice si es muestra o población, pero sé que no están todas las ciudades, entonces:

$$s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

$$s = \sqrt{\frac{1}{29} \sum (x_i - 8{,}75)^2}$$

$$s = \sqrt{\frac{1}{29} \left((0.6 - 8{,}75)^2 + (0.7 - 8{,}75)^2 + \dots \right.}$$

$$\left. \dots (103.5 - 8{,}75)^2 \right)$$

$s \approx 19.91$

1.2.) Voy a realizar un boxplot focalizado en la misma columna anterior (GDP (US$ billion))

⇒ Ya tengo los datos ordenados en el paso anterior

⇒ Q_1 = Mediana de la primera mitad de dato

Q_2 = Mediana

Q_3 = Mediana de la segunda mi

Q_1 = 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.5, 1.7, 1.8, 2, 2.1, 2.3

$$Q_1 = \frac{1,2 + 1,3}{2} = 1,25$$

Q_3 = 3, 3.2, 3.5, 3.8, 4, 4.8, 5.1, 6.2, 7.3, 10.5, 16.8, 22.4, 44.1, 103.5

$$Q_3 = \frac{5,1 + 6,2}{2} = 5,65$$

Q_2 = 2,65

X_{min} = ~~[illegible]~~ = 0,6
X_{max} = ~~[illegible]~~ = 103.5

limite Inferior =

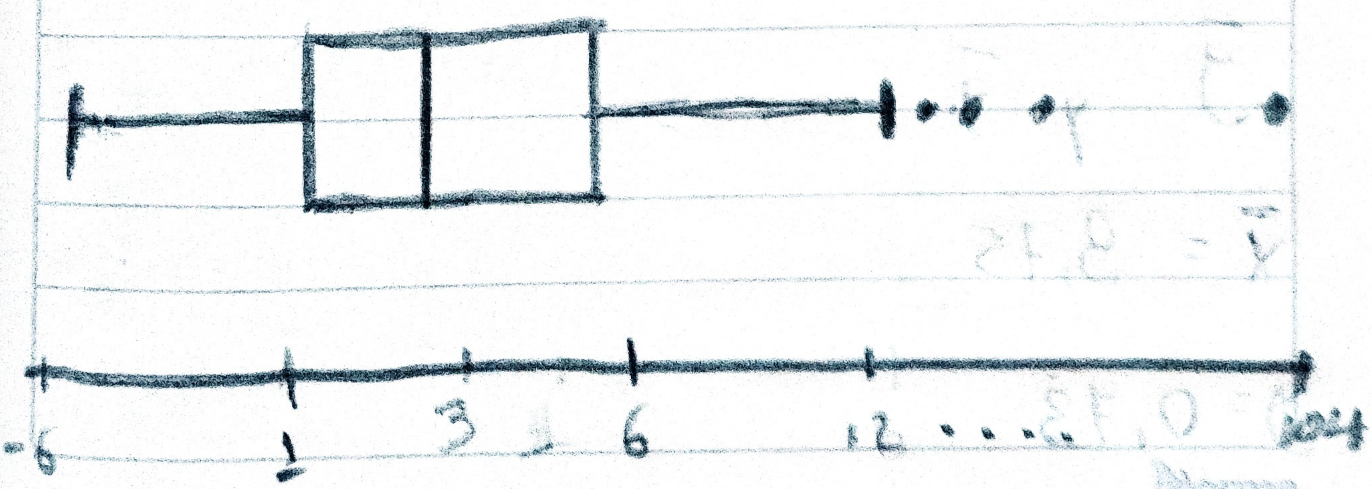
$$Q_1 - 1,5 * IQR = -5,35$$

limite Superior =

$$Q_3 + 1,5 * IQR = 12,25$$

$$IQR = Q_3 - Q_1 = 4;4$$

Aunqe no es perfecto va simular el boxplot



1.4) Quiero correlacionar GDP con population

→ Ya tengo $s, \bar{x}$ de GDP y sera

y partiendo de qe population sera

$$\Rightarrow \bar{y} = \frac{\sum y_i}{n}$$

$$s_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$$

pero sob necesitaremos

$\bar{y}$ y $\bar{x}$

$\bar{x} = 8,75$

$\bar{y} = 0,73$

$$\Rightarrow \sum (x_i - \bar{x})^2 = (0{,}6 - 8{,}75)^2 + (0{,}7 - 8{,}75)^2$$

$$+ (0{,}8 - 8{,}75)^2 + \ldots + (103{,}5 - 8{,}75)^2$$

$$= 6{,}64 + 6{,}48 + 6{,}32 + \ldots + 8977{,}56$$

$$= 11500{,}955$$

$$\sum (y_i - \bar{y})^2 = 53.054$$

$(x_i - \bar{x})(y_i - y) =$ Partiendo de que

$$(x_i - \bar{x}) = [-8{,}15,\ -8{,}05,\ -7.95 \ldots,\ 94{,}75]$$

$$(y_i - y) = [-0{,}72, \ldots, 6{,}45]$$

$$\Rightarrow \sum (x_i - \bar{x})(y_i - y) = 774{,}366$$

$\sqrt{a} \cdot \sqrt{b} = \sqrt{ab}$

$\rightarrow \dfrac{774{,}366}{\sqrt{11500{,}955 \cdot 53.054}}$

$= 0{,}9913$

Nos dice que aumentan siempre cuando la otra aumente (es la tendencia) y que es una relación fuerte