**Anexo 1. Detección de outliers**

La detección de *outliers* se realizó mediante dos (2) técnicas, la primera visual con cajas de bigotes para cada variable y la segunda usando el z-core y un umbral (threshold).

Se consideraron como outliers o valores extremos aquellos que estuvieran por debajo del límite inferior o por encima del superior de acuerdo con la siguiente fórmula:

Donde:

: umbral inferior por debajo del cual los datos se consideran *outliers*

: umbral superior por encima del cual los datos se consideran *outliers*

Q1 (Primer cuartil): Es el valor que separa el 25% inferior de los datos del 75% superior

Q3 (Tercer cuartil): es el valor que separa el 75% inferior de los datos del 25% superior

IQR: Rango intercuartílico. Se calcula como la diferencia entre Q3 y Q1

El factor de 1.5 es una regla empírica comúnmente utilizada en la técnica del boxplot para definir un rango razonable de variación de los datos.

Asimismo, también se incluyeron como *outliers* los valores reportados por encima o por debajo de 3 veces el Z-Score, definido para el análisis como:

Donde:

: valor individual de la variable i

: es la media de los datos de la variable i

*:* es la desviación estándar de la variable i

Esto garantiza que cualquier registro con un Z-Score que sea mayor que 3 se considere un *outlier* porque está a más de 3 desviaciones estándar de la media.