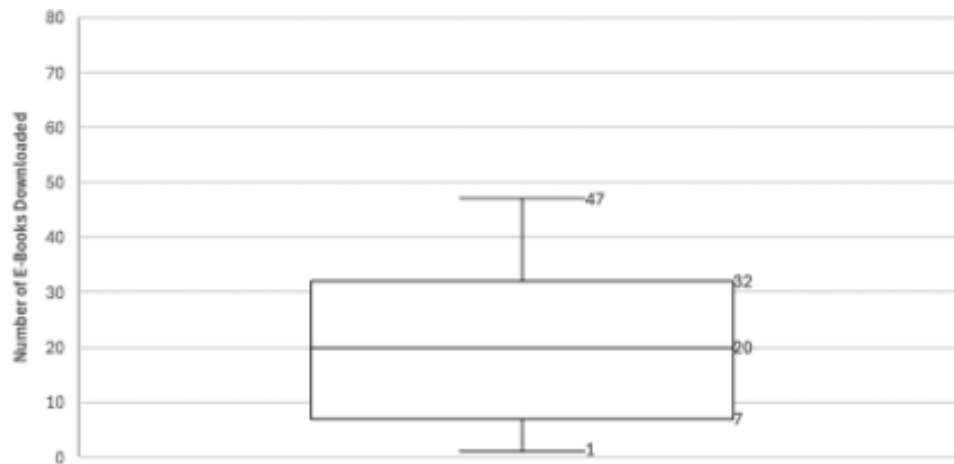Q1. What type of plot is most suitable to visualise and analyse the correlation between two variables? Why?

Q2. Given the boxplot below, calculate Q1, Q2, Q3, IQR, Upper Inner, Lower Inner Fence. Also give examples of outlier data points.

Q3. A researcher analyzes the relationship between age and income, where income increases rapidly with age up to a certain point and then plateaus. What would you expect if the researcher uses Pearson correlation instead of Spearman correlation.

#1 spearman

Q1. A scatterplot is the most suitable visualisation to observe and analyse the correlation between two variables because it shows the strength and direction of the correlation between two observed variables.

Q2.

   a. $Q1 = 7$

   b. $Q2 = \text{Median} = 20$

   c. $Q3 = 32$

   d. $\text{Max} = 47$

   e. $\text{Min} = 1$

   f. $\text{IQR} = Q3 - Q1 = 32 - 7 = 25$

   g. Upper Inner Fence $= Q3 + (1.5 * \text{IQR}) = 32 + (25 * 1.5) = 69.5$

   h. Lower Inner Fence $= Q1 - (1.5 * \text{IQR}) = 7 - (25 * 1.5) = -30.5$.

   i. Outlier data points would be any data points that are beyond either the lower inner fence or upper inner fence.

Q3.

Pearson correlation is the most common method for measuring the strength and direction of a linear relationship between two variables. It produces a value between -1 and 1, where 1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship. This means that when one variable increases or decreases, the Pearson correlation shows whether the other variable tends to increase or decrease in the same or opposite direction.

Spearman's correlation measures the strength and direction of the relationship between two variables when they are monotonically related. This means that the relationship is consistent in direction (either always increasing or always decreasing), regardless of whether it is linear or nonlinear. Spearman's correlation is particularly suitable for ordinal data, where the data can be ranked but the distances between ranks may not be equal. Like Pearson's correlation, Spearman's correlation produces a value between -1 and 1, capturing both linear and nonlinear monotonic relationships.