

Question :-

Explore the height of individual in a population. Calculate the range, quartile deviation, standard deviation, and variance using R.

Software Required :-

R software
R Studio

Description :-

The goal is to explore the distribution of heights within a population and calculate various statistical measures to understand its spread and variability. These measures include:

Range: The difference between the tallest and shortest individuals.

Quartile Deviation (IQR): The difference between the 75th and 25th percentiles, representing the range containing the middle 50% of the population.

Standard Deviation: A measure of how spread out the data is around the average height.

Variance: The square of the standard deviation, another measure of variability.

Procedure/Algorithm :-

Import the chosen data set or add in array.

Calculate the range using ***Range = Maximum value - Minimum value***.

Calculate quartiles, sort the data set in ascending order, **Q1** (first quartile) is the median of the first half of the sorted data. **Q3** (third quartile) is the median of the second half of the sorted data.

Calculate quartile deviation using ***quartile deviation = (Q3 - Q1) / 2***.

Calculate standard deviation, calculate the mean (average) of the data set. For each data point, subtract the mean and square the result.

Sum up all the squared differences. Divide the sum by the total number of data points – 1 and take the square root of the result.

Calculate ***Variance = (Standard deviation) ^2***.

Output the calculated range, quartile deviation, standard deviation, and variance

Code :-

```
heights <- c(156,182,170,165,178,160,175,68,172,178,155,181,186,147,148,149,150)

range_height <- diff(range(heights))

q1 <- quantile(heights, 0.25)
q3 <- quantile(heights, 0.75)
quartile_deviation <- (q3 - q1) / 2

std_deviation <- sd(heights)

variance <- var(heights)

hist(heights, main = "Height Distribution", xlab = "Height", ylab = "Frequency", col = "grey", border
= "black")

abline(v = mean(heights), col = "red", lwd = 2, lty = 2)
abline(v = median(heights), col = "green", lwd = 2, lty = 2)
abline(v = q1, col = "blue", lwd = 2, lty = 2)
abline(v = q3, col = "orange", lwd = 2, lty = 2)

legend("topleft", legend = c("Mean", "Median", "Q1", "Q3"), col = c("red", "green", "blue",
"orange"), lty = 2, lwd = 2)

cat("Range:", range_height, "\n")
cat("Quartile Deviation:", quartile_deviation, "\n")
cat("Standard Deviation:", std_deviation, "\n")
cat("Variance:", variance, "\n")
```

Values	
a	num [1:11] 32 5 8 16 12 8 6 20 8 15 ...
data	num [1:12] 5 2 7 3 5 8 1 6 4 5 ...
heights	num [1:17] 156 182 170 165 178 160 175 68 172 178 ...
q1	Named num 150
q3	Named num 178
qdev_value	1.75
quartile_deviation	Named num 14
range_height	118
range_value	num [1:2] 1 23
sd_value	6.17914380653325
std_deviation	27.1362303940691
variance	736.375

Output :-

```
R 4.3.2 · ~/
>
> hist(heights, main = "Height Distribution", xlab = "Height", ylab = "Frequency", col =
"grey", border = "black")
>
> abline(v = mean(heights), col = "red", lwd = 2, lty = 2)
> abline(v = median(heights), col = "green", lwd = 2, lty = 2)
> abline(v = q1, col = "blue", lwd = 2, lty = 2)
> abline(v = q3, col = "orange", lwd = 2, lty = 2)
>
> legend("topleft", legend = c("Mean", "Median", "Q1", "Q3"), col = c("red", "green", "blue", "orange"), lty = 2, lwd = 2)
>
> cat("Range:", range_height, "\n")
Range: 118
> cat("Quartile Deviation:", quartile_deviation, "\n")
Quartile Deviation: 14
> cat("Standard Deviation:", std_deviation, "\n")
Standard Deviation: 27.13623
> cat("Variance:", variance, "\n")
Variance: 736.375
> |
```

Visualization :-

